

ABSTRACT

Title of dissertation: **ESSAYS ON LABOR MARKETS AND
INFLATION**

 Edward Olivares, Doctor of Philosophy, 2023

Dissertation directed by: **Professor John Haltiwanger**
 Department of Economics

This dissertation presents research on several topics in economics.

Chapter 1 and Chapter 2 explore the implications of geographic labor mobility in the context of the US labor market. These chapters share a common motivation based on statistics presented herein describing the geographic dimension of job switching behavior; primarily that a surprisingly high share of job-to-job flows within the US take place across metropolitan areas. Chapter 1 explores the microeconomic implications of this for workers through the lens of non-local outside options. In labor markets with frictions, outside options are a key determinant of workers' wages and firms' rents. When outside options improve, workers can benefit by switching jobs, but even those that remain in the same job can realize gains through leveraging their improved bargaining position to renegotiate wages. In this chapter, I study the geographic dimension of workers' outside options and effects on labor mobility and wages for job stayers. I show that a large share of job-to-job flows are across metropolitan areas (MSAs), which suggests that the non-local dimension of labor market opportunities may be substantial. To obtain causal estimates of the effect of non-local out-

side options on wages and geographic labor mobility, I construct measures of exposure to changes in labor market conditions in other markets, and use a shift-share instrumental variable strategy to identify exogenous variation in non-local labor demand. I find that increases in labor demand in an MSA's network of labor markets are associated with increased job-to-job outflows with an elasticity of about .30, and higher wage growth for job stayers with an elasticity of .11. The effect of non-local shocks on job switching and wage growth is 30-50% of similar estimates of the effect of local labor demand shocks. Labor mobility is much more responsive to demand from MSAs with the strongest historic labor flows, which account for about 70% of the total mobility effect. I find similar mobility responses across education levels, but the effect of non-local outside options on wage growth is concentrated on workers without a 4-year college degree and in industries with lower average education levels.

Chapter 2 turns to the macroeconomic implications of geographic labor mobility in determining long-run labor market outcomes. Using U.S. data, I show that job-to-job flows across metro areas are about 40% of all metro area job-to-job flows, and that there is substantial heterogeneity across metro areas in the rate of incoming and outgoing job-to-job flows. I introduce a general equilibrium model of spatial on-the-job search that provides a framework for studying the effects of labor markets' heterogeneous geographic positions on long-run outcomes. In the model, labor demand is endogenous and wage bargaining allows me to explore the implications of worker's labor mobility on employer market power. I calibrate a simple version of the model and find that relative to an economy with no mobility, there is a moderate increase in average wages and a fall in unemployment. Changes in firm rents due to workers' stronger bargaining position account for about half of the increase in average wages.

In Chapter 3, I present coauthored work studying methods for the measurement of inflation

using large, micro-level retail sales data sources. In particular, this chapter explores alternative methods for adjusting price indices for quality change at scale. These methods can be applied to large-scale item-level transactions data that includes information on prices, quantities, and item attributes. The hedonic methods can take into account the changing valuations of both observable and unobservable characteristics in the presence of product turnover. This chapter also considers demand-based approaches that take into account changing product quality from product turnover and changing appeal of continuing products. This chapter provides evidence of substantial quality-adjustment in prices for a wide range of goods, including both high-tech consumer products and food products.

ESSAYS ON LABOR MARKETS AND INFLATION

by

Edward Olivares

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor John Haltiwanger, Chair/Advisor
Professor John Shea
Professor Judith Hellerstein
Assistant Professor Thomas Drechsel
Associate Professor Evan Starr

Preface

Chapter 3 is based on the working paper "Quality Adjustment at Scale: Hedonic vs. Exact Demand-Based Price Indices," coauthored with Gabriel Ehrlich, John Haltiwanger, Ron Jarmin, David Johnson, Luke Pardue, Matthew Shapiro, and Laura Yi Zhao. This research is representative of the output of a broader project studying the potential that modern data sources have for improvements in the measurement of inflation. All material is presented with permission from coauthors.

My individual contribution to this work includes development of hedonic models at the product group level for data from the NPD Group; descriptive analysis of product entry, lifecycle dynamics, and the distribution of product quality at the product group level for data from the NPD Group; methodological contributions including development of estimation methods for nested elasticity of substitution that build on methods introduced by [Hottman et al. \(2016\)](#) as well as as estimation methods for hedonic models including weighting and other specification details; and theoretical analysis of consumer demand shocks based on a constant elasticity of substitution model and the consumer valuation bias building on work by [Redding and Weinstein \(2018\)](#).

Acknowledgments

I am indebted to my advisor, John Haltiwanger, for his confidence, support, and knowledgeable counsel. I am also grateful to Judy Hellerstein and John Shea for their invaluable advice and suggestions. I wish to thank the co-authors of chapter 3 of this dissertation, Gabriel Ehrlich, John Haltiwanger, Ron Jarmin, David Johnson, Luke Pardue, Matthew Shapiro, and Laura Yi Zhao.

I acknowledge the financial support from the University of Maryland and the US Census Bureau's Economic Measurement Research Internship program.

I have greatly benefited from the camaraderie and commiseration among my fellow graduate students in the Economics department, and I particularly wish to thank Luke Pardue, Eugene Oue, Nathalie Gonzalez, and Chris Roudiez for their friendship and advice. Lastly, this dissertation was made possible through the boundless support of my wife, Larkin Willis.

Table of Contents

Preface		ii
Acknowledgements		iii
Table of Contents		iv
1	Labor Mobility, Nonlocal Outside Options, and Wages	1
1.1	Introduction	1
1.2	Related Literature	4
1.3	Motivating Evidence: Job-to-Job Flows and Geographic Labor Mobility	8
1.3.1	Data	8
1.3.2	Geographic Labor Mobility and Job Switching	11
1.4	Nonlocal Labor Demand, Labor Mobility, and Earnings Growth	16
1.4.1	Nonlocal Outside Options and the Job Finding Rate	17
1.4.2	Measures of Nonlocal Hires Growth	19
1.4.3	Empirical Framework	21
1.4.4	Network Instruments	22
	Local Shift-Share Instruments	23
	Network Instrument	23
	Identification	25
1.4.5	Estimating Equations	26
1.5	Results	29
1.5.1	Nonlocal Demand and J2J Outflows	29
	Dynamic Response of J2J Outflows	33
	Heterogeneous Effects	34
1.5.2	Nonlocal Labor Demand and Wage Growth	38
	Heterogeneous Effects	42
1.6	Conclusion	45
1.7	Appendix to Chapter 1	48
2	Geographic Labor Mobility and On-the-Job Search	55

2.1	Introduction	55
2.2	Related Literature	58
2.3	Motivating Facts	62
2.3.1	Data: Job-to-Job Origin-Destination Statistics	62
2.3.2	Sample	64
2.3.3	Descriptive Analysis	64
2.3.4	How important are Cross-Metro Job-to-Job Flows? Geographic Dispersion	65 66
2.3.5	MSA-level Heterogeneity	67
2.3.6	Recap	71
2.4	Model: Spatial On-the-Job Search	72
2.4.1	Environment	72
	Workers	72
	Firms	73
	Labor Market	73
2.4.2	Spatial Matching Function	73
2.4.3	Workers	75
	Labor Mobility: Job-to-job Quits	79
	Labor Mobility: Job-to-Job Hires	79
	Labor Mobility: Moves to Unemployment	80
2.4.4	Firms	80
2.4.5	Wage Bargaining	82
2.4.6	Steady State	85
2.4.7	Equilibrium	86
	Algorithm	87
2.5	Results	88
2.5.1	Environment and Parameters	88
2.5.2	Mobility-Compatible Indifference Wages	91
2.5.3	Wages and Markdowns	94
2.5.4	Counterfactuals	95
2.6	Conclusion	98
2.7	Appendix to Chapter 2	101
2.7.1	Firms' Value Functions	101
3	Quality Adjustment at Scale: Hedonic vs. Exact Demand-Based Price Indices	103
3.1	Introduction	103
3.2	Data	106
3.2.1	NPD Data	107
3.2.2	Nielsen Data	108
3.3	Conceptual Framework	109
3.3.1	Traditional Price Indices	109
3.3.2	Hedonic Price Indices	110

3.3.3	Demand-Based Price Indices	115
3.4	Results	119
3.4.1	NPD Results	119
	Hedonics	119
	CES Demand-Based Price Indices	121
	Comparing Traditional, Hedonic, and Exact Price Indices	125
3.4.2	Nielsen Results	126
3.4.3	Chain Drift	129
3.4.4	Taking Stock	130
3.5	Conclusion	133
3.6	Appendix to Chapter 3	148
3.7	Additional Price Index Details	148
3.7.1	Traditional Price Indices	148
3.7.2	Hedonic Estimation Details	149
3.8	Using the Nielsen Data	152
3.8.1	Nielsen Data Preparation	152
3.8.2	Comparisons of the Nielsen Data to Official Statistics	153
3.8.3	Product Descriptions in the Nielsen Data	155
3.8.4	Common Goods Rules – Consumer Panel and Retail Scanner	155
3.8.5	Machine Learning and Hedonics	157
3.8.6	Assessing the National Market Assumption in the Nielsen Data	159
3.9	Additional Evidence on the Behavior of the CUPI	160
3.9.1	Behavior of the CUPI with a Nested Preference Structure	161
3.9.2	Analytical Characterization of the Taste Shock Bias	162
3.9.3	Simulation Evidence on the Behavior of the CES Exact Price Indices	165
	Simulation Model Environment	165
	Simulation Evidence	168
3.9.4	Appendix Tables and Figures	174
4	Conclusion	197
	Bibliography	199

List of Tables

1.1	Job-to-Job Flow Rates Across MSAs by Education Level	11
1.2	Share of Cross-Metro Job-to-Job Flows Within	15
1.3	Nonlocal Demand and J2J Outflows	30
1.4	Labor Demand and Local Job-to-Job Flows	32
1.5	Dynamic Response of J2J Outflows to Nonlocal Demand: Net- network IV Estimates	34
1.6	Nonlocal Demand and J2J Outflows: IV Estimates, Closely Con- nected Destinations	36
1.7	Nonlocal Demand and J2J Outflows: Effect by Education Level	37
1.8	Nonlocal Demand and J2J Outflows: Annualized Estimates	39
1.9	Nonlocal Demand and Earnings Growth	40
1.10	Local Demand Shocks and Wage Growth: Bartik IV	41
1.11	Nonlocal Demand and Wage Growth: IV Estimates, Closely Con- nected Destinations	43
1.12	Nonlocal Demand and Earnings Growth: Heterogeneous Effects by Industry and Education Level	44
A1.1	MSA Sample Summary Statistics	48
A1.2	CPS Sample Summary Statistics	49
A1.3	MSA-Level Observations in CPS Sample	49
A1.4	First Stage Network IV Results	50
A1.5	Robustness Check: Nonlocal Demand and Local J2J Flows, Top Destinations	50
A1.6	Nonlocal Demand and Earnings Growth: Excluding Low-Observation MSAs	51
A1.7	Nonlocal Demand and Wage Growth: Nominal vs. Real Wage Growth Effects	51
A1.8	Nonlocal Demand and Earnings Growth: Job Stayers + Local Job Switchers	52
A1.9	Nonlocal Demand and J2J Outflows: No Local Controls	52
A1.10	Nonlocal Demand and Wage Growth: No Local Controls	53
A1.11	Nonlocal Demand and Wage Growth: No Recent Job Switches and Same Industry/Occupation Observations	54
2.1	Within-Metro and Across-Metro Job-to-Job Flow Rates	65

2.2	Share of Cross-Metro Job-to-Job Flows Within:	67
2.3	J2J Outflows as a share of all J2J Separations	68
2.4	Gravity Regression, J2J Flows to Destination Metro	70
2.5	Gravity Regression, J2J Flows to Destination Metro	71
2.6	Parameters	91
2.7	Calibration: Targeted Model Moments	92
2.8	Comparative Statics	97
2.9	Decomposing Change in Average Wages	98
3.1	Rates of Product Turnover: NPD Data	136
3.2	Estimated Elasticities of Substitution: NPD Data	136
3.3	Alternative Price Indices, Levels in 2018q4 Relative to 2014q4: NPD Data	136
3.4	Alternative Price Change Indices, Chained (C) vs GEKS-Lite (GL), NPD Data	137
3.5	Alternative Price Change Indices, Chained vs GEKS-Lite, Nielsen, Food	137
3.A.1	Impact of Alternative Imputation for Missing Price of Entrants, NPD Data	174
3.A.2	Comparison of GEKS-Lite and Rolling Year GEKS, Annual Chained Price Indices, NPD Data	174
3.A.3	Summary Statistics for Alternative Price Chained Indices, GEKS- Lite, NPD Data	174
3.A.4	Hedonic Models: Goodness of Fit	175
3.A.5	Nested Estimated Elasticities of Substitution: NPD Data	175
3.A.6	Food Product Groups: Nielsen Retail Scanner Data	176
3.A.7	Nonfood Product Groups: Nielsen Retail Scanner Data	176

List of Figures

1.1	Cross-MSA Job-to-Job Flows: Trends Over Time	13
1.2	Cross-MSA Job-to-Job Flows: MSA-Level Variation	16
2.1	Within-Metro and Cross-Metro J2J Rate	66
2.2	Mobility-Compatible Indifference Wage Premium	93
2.3	Relative Labor Mobility along the Job Ladder	94
2.4	Distribution of Productivity and Wages	95
2.5	Markdown of Wages from Productivity	96
3.1	Product Lifecycle Dynamics	138
3.2	Hedonic Specifications: Fixed vs. Time-Varying Unobservables	139
3.3	Hedonic Specifications: Evaluating Time-Varying Unobservable Specification	140
3.4	Components of Feenstra and UPI	141
3.5	CUPI: Alternative Common Goods Rules (CGRs)	142
3.6	Comparison of Main Price Index Specifications	143
3.7	Main Price Index Specifications, Without CUPI	144
3.8	Main Price Index Specifications: Cumulative Price Level Changes	145
3.9	CUPI and Its Components with Alternative CGRs, Nielsen Food	146
3.10	Main Price Index Specifications: Price Changes and Levels, Nielsen Food	147
3.A.1	Alternative Hedonic Estimation Strategies NPD Data	177
3.A.2	Test of Time-Varying Unobservable Hedonic Specification, First- Differences and Levels Estimation	178
3.A.3	Main Price Index Specifications: Cumulative Price Level Changes, No CUPI	179
3.A.4	Nested CUPI: Characteristics- and P-Hat- Based Nests Percent Changes . .	180
3.A.5	Dispersion of Relative Product Appeal	181
3.A.6	Sales-weighted and Unweighted Distributions of Market Penetra- tion of Items in Nielsen Data, Food	182
3.A.7	PCE vs Nielsen Sales for Scanner and Consumer Panel, Food and Nonfood .	183
3.A.8	PCE vs Nielsen Sales for Scanner, By Category within Food and Nonfood .	184
3.A.9	BLS CPI vs Nielsen Laspeyres, Food and Nonfood	185
3.A.10	Common Goods Rules–2-quarter vs 5-quarter Horizons	186

3.A.11	Common Goods Rules – Nielsen Consumer Panel	187
3.A.12	Common Goods Rules – Nielsen Consumer Panel	188
3.A.13	Replication of Redding and Weinstein (2020) with Nielsen Consumer Panel	189
3.A.14	Common Goods Rules – Nielsen Scanner Panel	190
3.A.15	Main Price Index Specifications: Price Changes and Levels, Nielsen Nonfood	191
3.A.16	Simulated CES Exact Price Indices with Trends in Cost Shifters	192
3.A.17	Simulated CES Exact Price Indices with Trends in Dispersion of Product Appeal	193
3.A.18	Simulated CES Exact Price Indices with Product Turnover	194
3.A.19	Simulated CES Exact Price Indices with Segmented Markets	195
3.A.20	Simulated CES Exact Price Indices with Partial Stock-outs prior to Exit	196

Chapter 1: Labor Mobility, Nonlocal Outside Options, and Wages

1.1 Introduction

Job switching contributes to earnings growth for individual workers and is a source of productivity-enhancing reallocation that contributes to aggregate growth (Topel and Ward, 1992; Haltiwanger et al., 2017). In addition, workers that receive outside job offers can leverage their outside options in wage negotiations even without switching jobs, reducing employer market power and markdowns of wages from productivity (Cahuc et al., 2006; Beaudry et al., 2014; Caldwell and Harmon, 2019). While labor markets have a strong geographic element, because of migration, commuting, and (increasingly) remote work, labor markets are not strictly local. How do workers respond to changes in job opportunities in other labor markets through job switching and wage renegotiation?

In this paper, I study nonlocal outside options and the transmission of labor demand across regions through the lens of job-to-job (J2J) flows across US metropolitan areas (MSAs). I show that the nonlocal dimension of workers' outside options – reflected by J2J flows – is quite substantial: about 40% of J2J flows are across, rather than within MSAs.¹ I then provide causal estimates of for the effects of changes in nonlocal outside options on reallocation of workers through increased J2J outflows, as well as greater wage growth for workers.

Intuitively, job-to-job flows are a natural indicator of workers' outside options: when a worker switches jobs from one firm to another, this indicates a better job opportunity (or in

¹While some of these flows are surely due to people living on the margin between different MSAs, about half of those flows are across states, and even more are across commuting zones. Moreover, the average distance between city centers is over 100 miles, suggesting that the non-local dimension of job flows is not simply a product of insufficient geographic aggregation.

the case of involuntary separations, a next-best opportunity). At the city level, job-to-job outflows are a strong indicator of the opportunities available to workers outside of their local labor market.

To evaluate the contribution of nonlocal job opportunities to labor market outcomes, I develop measures of changes in non-local outside options and estimate their effect on 1) geographic labor mobility, and 2) earnings growth. The measures of non-local outside options are based on exposure through historic job-to-job flow networks to growth in hires in other MSAs. Intuitively, stronger hiring growth in other labor markets increases the likelihood of a job searcher receiving a nonlocal outside offer, which the worker can exercise by switching jobs or leverage to renegotiate wages at their current job. I formalize this idea by showing that the growth rate of hires is a proxy for the job finding rate, though an imperfect one as hiring also reflects changes in labor supply. However, it is plausible to assume that labor demand shocks only influence hiring growth through the job finding rate, which motivates the instrumental variable strategy based on the Network IV specification of [Schubert \(2021\)](#). In this approach, shift-share instruments are used to identify exogenous variation in labor demand in all other MSAs, rather than in the local market.

I find that non-local labor demand shocks induce increased J2J outflows with an elasticity of roughly 0.3. The response of J2J outflows is strongest in the initial quarter but shows some persistence: the cumulative increase in J2J outflows after four quarters is substantially higher at about 0.56. While this elasticity is large in relative terms, the identified shocks to nonlocal labor demand are fairly small, so their aggregate contribution is limited. I find that a typical shock to nonlocal demand leads to an increase in J2J outflows of about 1.72% in the short run (about 30% of the effect size of a similarly estimated response to a local labor demand shock) and 3.25% over the following year, or a .025 – .046 percentage point increase for an MSA with the average quarterly J2J outflow rate of 1.41%. Furthermore, these estimates imply that demand-driven reallocation of workers accounts for no more than 10% of overall cross-MSA job switches, which suggests a larger role for other factors such as local labor market conditions, individual labor supply factors, and idiosyncratic variation in individual job opportunities that is not associated with aggregate shocks.

Despite the muted response of labor mobility to nonlocal demand in absolute terms, there

is heterogeneity in the response of mobility to nonlocal demand that points to a larger role for certain types of destinations or workers. I find that J2J outflows are about 30% more responsive to demand shocks from MSAs that have overlapping commuting zones with the origin (though note that the benchmark estimate is robust to limiting the sample to strictly cross-commuting zone J2J flows). This is intuitive, as the option to commute rather than migrate reduces the fixed costs involved with exercising a nonlocal outside option. Furthermore, I find that the majority of the response of J2J outflows to nonlocal demand can be accounted for by demand from a handful of MSA-specific top destinations; while the elasticity of J2J outflows with respect to nonlocal labor demand shocks from top destinations is similar in magnitude to the benchmark estimate, these top destinations account for a large share of the variation in nonlocal demand due to the skewed nature of cross-MSA job switches. Lastly, I find some evidence of heterogeneity in mobility by education level, with the strongest evidence pointing towards workers without a high school degree being significantly less responsive to nonlocal labor demand.

I then turn to estimating the effect of nonlocal labor demand on wage growth. Importantly, I limit my analysis to workers that report no recent job switches. This isolates the variation in outside options independent of worker and firm effects and changes in the composition of employment. I find that higher nonlocal labor demand causes increased wage growth for job stayers with an elasticity between .11 and .15. A typical shock to nonlocal outside options leads to a 1.24%-1.73% increase in wages. This is about 24% of the effect of a similarly estimated response of wage growth to local labor demand, which leads to a 5.2% increase in wages. This difference is driven both by a higher elasticity of wages with respect to local demand, and the relative size of local and nonlocal labor demand shocks.

I find less evidence of heterogeneity in the response of wages to nonlocal demand with respect to destination characteristics: shocks from closely connected destinations have similarly sized effects as other destinations. I find substantial heterogeneity in the wage growth effect by education level and industry. The strongest effect of nonlocal demand on wages is for workers with some college or an associate's degree, while the effect for workers with a college degree or greater is much smaller and statistically null. This aligns with the effects by industry of employment, which are strongest in local service industries, such as

retail and hospitality, as well as manual industries such as manufacturing, construction, and transportation. With the exception of education and healthcare, industries with the highest level of education, such as skilled scalable services (Eckert et al., 2020), tend to have the lowest estimated elasticity of wages with respect to nonlocal labor demand.

In the following section I provide a review of related literature on monopsony in labor markets and workers' outside options, regional adjustment and the effects of local labor demand, and the consequences of migration and other forms of labor mobility. In section 3, I present descriptive statistics on the importance of cross-MSA job-to-job flows in the US. In section 4, I outline my empirical strategy for estimating the effect of nonlocal outside options on labor mobility and wages, including measures of exposure to nonlocal labor market conditions and the instrumental variables approach I use to isolate variation in nonlocal demand. In section 5 I present estimation results, and section 6 concludes.

1.2 Related Literature

This paper studies the geographic dimension of workers' labor market opportunities, the role of outside options in wage determination, and the responsiveness of labor mobility and wages to nonlocal labor demand. This builds on several strands of economic research.

First is the literature that studies outside options in the labor market and their role in wage bargaining, labor mobility, and employer market power, which this paper contributes to by detailing the geographic dimension of outside options. One major contribution in this literature is Jager et al. (2019), which uses reforms to unemployment insurance (UI) in Austria to study the effects of changes in the value of non-employment on wages for the employed. They find minimal effects of increases in UI benefits on wages: less than \$0.01 per \$1.00 of benefit increase. A number of other studies use variation in the value of or likelihood of receiving an outside job offer, which may be the more relevant measure of outside options for currently employed workers. For example, Caldwell and Harmon (2019) uses Danish administrative data to show that changes in worker-level outside options predict both job switches and earnings growth. The empirical strategy relies the fact that networks of former coworkers are a channel of information transmission about labor market opportunities: they find that increases in hiring at the firms of former coworkers are associated with wage gains

and increased job-to-job transitions. [Lachowska et al. \(2021\)](#) similarly studies the effect of worker-level outside options, using growth in coworkers' wages at the secondary jobs of dual job holders as an indicator of change in workers' outside option, finding that wage increases at secondary jobs increase both separation probabilities and wages at primary jobs. Both [Caldwell and Harmon \(2019\)](#) and [Lachowska et al. \(2021\)](#) find wage elasticities for job stayers that are relatively small (about .02), in large part due to the infrequency of renegotiating wages. Lastly, [Caldwell and Danieli \(2018\)](#) use dispersion of different types of workers across different types of jobs (including geographic dispersion) as a proxy for outside options. Using administrative data from Germany and instruments based on the introduction of high-speed rail and a shift-share ("Bartik") approach, they find that higher outside options are associated with higher wages. Their estimates of the elasticity of wages with respect to outside options range from .17 to .32, which is substantially greater than estimates from the previously mentioned studies using worker-level variation in outside options.

Relatedly, a number of studies analyze imperfect competition in local labor markets through the lens of employment concentration, in some cases accounting for labor mobility as a mediating factor. [Benmelech et al. \(2018\)](#), [Azar et al. \(2017\)](#), and [Rinz \(2018\)](#) estimate the relationship between local concentration and wage rates in the US, finding that greater concentration is associated with lower wages as well as a secular decline in local employment concentration in recent decades. [Marinescu et al. \(2019\)](#) applies the IV approach of [Azar et al. \(2017\)](#) and [Rinz \(2018\)](#) to concentration of new hires, rather than employment, using French administrative data, finding that greater concentration leads to reduced hires and wages. [Berger et al. \(2022\)](#) estimate a market-share dependent response of wages and labor supply with respect to within-firm variation in changes to state corporate tax rates, and combined with an oligopsonistic model of labor markets, find considerable welfare losses from employer market power. [Azar et al. \(2019\)](#) estimates employer market power within geography-occupations bins using a model of job differentiation and IV estimates of labor supply elasticities, finding average markdowns of wages from productivity of about 17%. [Arnold \(2019\)](#) estimates a negative effect of local employment concentration on earnings using mergers for identification and accounting for cross-industry mobility. Lastly, [Schubert et al. \(2020\)](#) uses a shift-share research design to estimate the wage effect of workers'

outside options using within-occupation employment concentration and cross-occupation mobility, finding an semi-elasticity of wages with respect to outside options between .24 and .37.

I contribute to the literature on employer market power and outside options in several ways. First, I isolate variation in the geographic dimension of workers' outside options by studying the effect of nonlocal labor demand on geographic labor mobility and wages. My estimates indicate that nonlocal outside options play an important role in workers' overall labor market opportunities, and influence both job switching and wages. Additionally, the findings in this paper that nonlocal outside options can effect job switching and wages demonstrate that in addition to cross-industry and cross-occupational mobility, geographic labor mobility is an important mediator of the effect of local labor market conditions on employer market power and other labor market outcomes. This has important implication for research on monopsony, as there is substantial heterogeneity in MSA-level job-to-job outflows, which suggests that in markets with limited local job opportunities where there is greater potential for monopsony, workers may respond by widening their job search.

Another related literature studying the effects of shocks to local labor markets. This includes the seminal work [Bartik \(1991\)](#) introduces the shift-share measures which interact local industry employment shares with national shocks to industry growth, which have been used widely since as instruments for local demand, and finds that local labor demand shocks have persistent effects on labor force participation and wages. [Blanchard and Katz \(1992\)](#) takes a VAR approach to studying regional adjustment in response to local labor demand shocks, finding similar short-run effects of local labor demand shocks, but that net migration is an important margin of adjustment and that causes the effects of local shocks to be short-lived. Further work including [Bartik \(1996\)](#) and [Bartik \(2015\)](#) has studied the distributional effects and state-dependence of local labor demand shocks, finding that initial unemployment rates and employment growth rates are an important mediator of the short and long run responses. [Dao et al. \(2017\)](#) applies the methods of [Bartik \(1991\)](#) and [Blanchard and Katz \(1992\)](#) to study the evolution of regional adjustment over time, finding that cross-state mobility in response to labor market shocks has fallen since the period studied in [Blanchard and Katz \(1992\)](#) covering the 1970s and 1980s, leading to more per-

sistent effects of local shocks. They do find, however, that over the sample period I study (post-2003), the responsiveness has increased somewhat, largely due to increased mobility in response to positive local shocks during the Great Recession. I contribute to this literature in two ways. First, this literature demonstrates the spillover of local labor demand shocks through a labor mobility channel, and I build on this by estimating spillover directly via the effects of non-local labor demand shocks. I am also able to demonstrate an additional mechanism that local demand shocks affect other markets by estimating the effects on outside options and wages for non-movers. Second, the research design I use applies the Network Bartik IV approach used by [Schubert \(2021\)](#) to study spillover of housing demand across markets, which builds on the Bartik IV approach to studying local labor markets. I apply this to the effects of labor demand shocks and in particular the effects of differential exposure to non-local labor demand shocks based on historic cross-MSA labor mobility.

Lastly, this paper contributes to a literature studying the labor market and macroeconomic consequences of migration and other forms of labor mobility. A number of papers, including [Kaplan and Schulhofer-Wohl \(2017\)](#), [Molloy et al. \(2016\)](#), [Molloy and Smith \(2019\)](#), [Kaplan and Schulhofer-Wohl \(2017\)](#), and [Hyatt et al. \(2018\)](#) have studied trends in US migration and geographic labor mobility, which have declined in recent decades. The link between job switching and migration has been demonstrated empirically by [Molloy et al. \(2014\)](#) and [Molloy et al. \(2017\)](#), and the consequences of geographic mobility frictions in the context of job search models has been studied by [Beaudry et al. \(2014\)](#), [Schmutz and Sidibé \(2018\)](#), and [Heise and Porzio \(2019\)](#), among others. Other work, such as [Monte et al. \(2018\)](#) and [Flemming \(2020\)](#) studies the role of commuting in the response to localized labor market shocks and in workers' job switches. Lastly, work including [Manning and Petrongolo \(2017\)](#) and [Marinescu and Rathelot \(2018\)](#) studies the role of geography in workers' job search behavior, finding that workers are highly sensitive to geographic proximity in their job searches. I contribute to this literature by providing descriptive analysis of job-to-job flows across U.S. metropolitan areas, which comprise a substantial share of all U.S. job switches, and display less of a secular trend than migration rates. I find that this aspect of geographic mobility, which is directly linked to labor mobility, has a wide degree of geographic dispersion in the aggregate: there are a substantial number of longer distance job-to-job flows; though the effects of nonlocal demand are more localized in that mobility

is more responsive to demand from nearby MSAs and those that have strong historic labor flows. Additionally, with my estimates I am able to evaluate the extent to which geographic labor mobility is driven by labor demand or labor supply. I find that though the share of geographic labor mobility driven by labor demand is small, accounting for as much as 10% of overall mobility.

1.3 Motivating Evidence: Job-to-Job Flows and Geographic Labor Mobility

In this section, I provide descriptive evidence on geographic mobility in the US labor market using data on quarterly cross-metro job-to-job flows that will motivate the empirical analysis of spillover of localized labor demand shocks across geographically distinct labor markets.

1.3.1 Data

For data on labor mobility across U.S. cities, I use the Job-to-Job Origin-Destination statistics (J2JOD) from the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) database. The LEHD is a matched employer-employee dataset constructed from administrative records from state level unemployment insurance agencies linked to establishment level data from the Quarterly Census of Employment and Wages (QCEW). These records cover about 95% of private sector employment (Hyatt et al., 2015) and are used to construct employer and employee work histories. The J2JOD statistics are a public use dataset that show inter-state, metro area, and industry job-to-job hires starting in the second quarter of 2000. The data are also available by detailed worker and firm characteristics. For workers, this includes education, age, race, and ethnicity. For firms, statistics are available by size and age categories. For the analysis in this paper, I use MSA and MSA-by-education level data on job-to-job transitions.

A job-to-job flow in the J2JOD is defined as a hire of a worker that recently was employed at another firm. Because workers can have multiple sources of earnings within each quarter, job-to-job flows are constructed with respect to a worker's dominant job, or the job that has the highest combined earnings in the current and previous period. For example, if a worker

is hired into a new job during period t and then works only at this job for a full period, then in period $t + 1$ this new job will be coded as the dominant job. Since the worker has earnings in the new job in both t and $t + 1$, then the worker was presumably employed in this job at the end of period t . If the worker had a dominant job coded at another firm at the beginning of period t , then this hire will be coded as a job-to-job flow in period t ². The J2JOD lists quarterly counts of job-to-job flows by origin MSA and destination MSA. If the original job and new job are in different MSAs, then the J2J switch will count as a J2J outflow for the origin and a J2J inflow for the destination.

Job-to-job flows can be further broken down into stable flows and non-stable flows. A stable job is defined as employment that lasts for at least one full quarter. For this to be the case, employment must be observed for at least three consecutive quarters. A stable job-to-job flow occurs when a worker transitions from one stable job to another.

The second source of labor market data is the Quarterly Workforce Indicators (QWI), which are also constructed from the LEHD, and report gross hires and separations, average earnings, and other labor market statistics for locations and industries. I use MSA, MSA-by-education level, and MSA-by-industry level employment, hires and earnings in the QWI.

The underlying data from the LEHD is constructed on a state-by-state basis, and not all states are available in each year. To ensure that statistics are consistent over time, I construct a balanced panel using metro areas that are available from 2003q1 to 2019q4. This preserves about 90% of the metro pair-quarter observations in the full sample, and the excluded metro areas are generally small. There are two notable exclusions: the Washington-Arlington-Alexandria metro area, which does not enter the LEHD until 2005, and the Boston-Cambridge-Newton metro area, which does not enter until 2010.

For worker-level outcomes, I use the Current Population Survey (CPS) Outgoing Rotation Group (also called the Earner Study) public-use microdata (CPS-IPUMS) over the period

²The J2JOD statistics also include job-to-job flows with a brief period of non-employment. The previous example was a within-quarter job switch because the worker was employed in the first job at the beginning of period t , and employed at the new job in end of period t . An adjacent-quarter job-to-job flow would take place if the worker was not employed in the first main job at the beginning of period t , but was in period $t - 1$. The J2J documentation recommends using the sum of both types of job-to-job flows, and I follow this convention.

2003-2019. The CPS interviews workers on a 4-on, 8-off, 4-on schedule. In 4th month of the initial and final interview periods, a subset of workers (about 20%) are asked additional questions about their earnings; this is the Outgoing Rotation Group. This provides up to two earnings observations per respondent in the ORG sample, which allows me to construct measures of year-over-year wage growth. To construct wages and wage growth, I largely follow the methodology of the Atlanta Fed's Wage Growth Tracker, which uses data from the CPS to produce monthly series of wage growth for different groups of workers. First, workers are asked to report their weekly earnings or hourly wages, depending on their compensation. For salaried workers, I convert weekly earnings to hourly wages by dividing by usual weekly hours or weekly hours in the previous week if usual hours are not reported. A substantial number of observations have imputed earnings or hourly wages, which I drop. I also drop any observations with topcoded earnings or wages (\$150,000 or the hourly equivalent) as they may bias wage growth calculations. For salaried workers, I further restrict the sample to workers that report absolute changes in hours of 10 or less (about 90% of the sample), which limits the effect of changes in hours on calculations of hourly wages. Lastly, I restrict the sample to prime-age workers (25-55) to remove those still in higher education or near retirement whose evaluation of job opportunities may be more influenced by geographic limitations and other specific job characteristics.

In order to more precisely identify the effect of nonlocal outside options on wages, I wish to limit the analysis to workers that remain in the same job in both wage observations³. However, in the CPS workers report job switches on a monthly basis, and are not asked in survey month 5 (the first month of the second survey wave) if they are at the same employer. This means that I do not observe job-staying or switching behavior in the 4-month gap between survey waves, so cannot completely exclude local job switches. In the benchmark sample, I restrict observations to those that report no job switches in the final month, i.e., immediately prior to the second wage observation. As a robustness check, I limit the sample to respondents that report no job switches in any month during the second survey wave and that report working in the same industry and occupation in both wage observations. This excludes about 16% of the benchmark sample.

³Note that only job switches are observed in the CPS, as workers that move either locally or to a different MSA will necessarily drop out of the survey, which uses a household frame.

1.3.2 Geographic Labor Mobility and Job Switching

Table 1.1 lists within-metro and across-metro job-to-job flow rates as a percentage of employment, hires, and all job-to-job flows. The average quarterly job-to-job flow rate is just over 4% of employment, 1.7% of which are across metropolitan areas. In the aggregate, cross-metro job-to-job hires make up about 40.3% percent of total job-to-job hires in metro areas, or 11.6% of total hires.⁴ Overall, this suggests an important role for interactions between geographically distinct labor markets in shaping workers’ job opportunities and firms’ recruitment pools.

Table 1.1: Job-to-Job Flow Rates Across MSAs by Education Level

	J2J Rate	Cross-Metro J2J Rate		% Across Due to:		
	% Emp	% Emp	% Hires	% J2J	Churn	Net Flows
All Jobs	4.3	1.7	11.6	40.3	94.6	5.40
Stable Jobs Only:						
All	2.7	1	13	39.1	92.6	7.4
Missing (Under 24)	5.6	2.2	11	39.2	91.1	8.9
Any College	2.3	0.9	16	41	92	8
Non-College	2.6	1.1	15	41.1	92.9	7.1

Average quarterly rates from 2003-2019. Stable jobs last for at least 3 quarters. Source: Job-to-Job statistics and Quarterly Workforce Indicators, U.S. Census Bureau. Own calculations.

In the lower panel of Table 1.1, I report the same statistics but restricting the sample to flows between stable jobs (those that observed in at least 3 quarters) and breaking out by education level. In my subsequent analysis of nonlocal labor demand, I use stable job-to-job flows to limit the influence of seasonal labor market or migration patterns. Almost by construction, stable jobs have a lower job-to-job flow rate, at 2.7% of employment. However, the share of job-to-job flows that is across MSAs is remarkably similar, at 39.1%. The share of cross-MSA J2J flows is also about 40% for workers with different education levels, though note that the overall job-to-job flow rates differ, with more educated workers switching jobs less frequently.

⁴Note that this is not the share of total job-to-job flows in the U.S. because the sample is restricted to job-to-job flows within or across metro areas. Job-to-job hires that originate from or end in a location outside of a metro area are not included in this count. Rather, this tells us the share of job-to-job flows starting and ending in a metro area that end in a metro different than the origin. Because this includes all within-metro flows but not all flows to or from each metro, this figure likely understates the extent of labor mobility that is relevant for metro areas.

A substantial share of cross-MSA job-to-job flows are accounted for by churn, rather than net migration. The last two columns of Table 1.1 decompose cross-metro job-to-job flows into net flows, defined as the sum of the absolute value of net cross-metro J2J hires for each metro, and excess flows, or the reallocation of workers across metro areas above what is needed to account for net reallocation. I find that this excess reallocation or churn accounts for most cross-metro job switches: net flows account for less than 10% of all cross-metro job switches. This is important to note, as I will use data on gross job-to-job outflows. Intuitively, if cross-MSA labor mobility was primarily driven by net flows, this would indicate underlying structural changes, and J2J outflows would be an indicator of local decline. Given the high degree of churn, J2J outflows are a sign of elevated opportunities available to workers, rather than stagnant or falling local labor supply.

To see how cross-metro job-to-job flows have evolved over time, Figure 1.1 plots the overall J2J rate as a percentage of employment, decomposed into within-metro and across-metro components. The total J2J rate is very cyclical, with a deep drop during the Great Recession. Since 2010, the J2J rate has recovered and in by 2017 surpassed its pre-recession peak. During the strong labor market of 2018 and 2019, job-to-job flows rose dramatically. The cross-metro J2J rate follows a similar pattern, with a dip during the recession and a surprisingly strong recovery. Cross-metro flows are less pro-cyclical than within-metro flows, so their share of total J2J actually rose slightly during the Great Recession, and fell during 2018-2019.

One notable feature of Figure 1.1 is that there is virtually no secular trend in cross-metro job-to-job flows from 2003 to 2018. This is puzzling, since previous literature has documented a decline in migration rates over recent decades (Kaplan and Schulhofer-Wohl, 2017; Molloy et al., 2016) and has linked this decline to declining labor reallocation (Molloy et al., 2017; Molloy and Smith, 2019). Based on this literature, one would expect to see cross-metro flows declining in importance. There are several factors that can explain this discrepancy without contradicting these previous findings. First, the secular decline in migration is a phenomenon dating back to the 1980s, and my sample only starts in 2003, so it is possible that despite the local peak in the late 2010s, the share of cross-metro J2J hires remains below levels observed in the 1990s or the earlier 2000s. Second, while

Figure 1.1: Cross-MSA Job-to-Job Flows: Trends Over Time

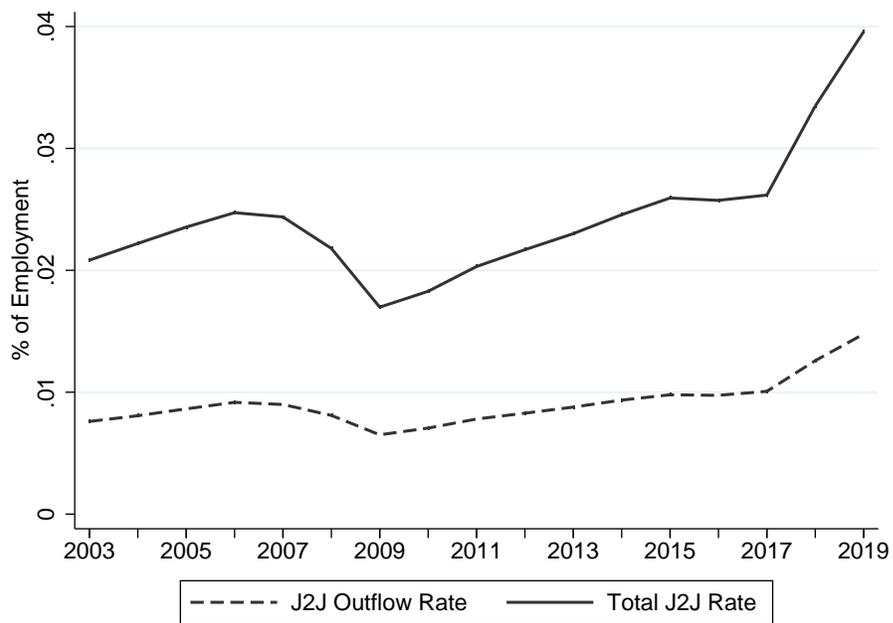


Figure 1.1: Average quarterly cross-MSA J2J outflow rate and overall J2J rate at the national level. Data: J2JOD, QWI

job-related mobility is an important component of migration (Molloy and Smith, 2019), job-to-job flows exclude moves of the unemployed, which may follow a separate trend. Some supportive evidence can be seen in the declining response of migration to regional shocks documented in Dao et al. (2017), which suggests that the unemployed have become less likely to move over time.

While cross-metro J2J hires are substantial, this does not on its own imply that spatial linkages between labor markets are important. One alternative explanation is errors in geographic aggregation: if labor markets are truly local but metro areas are smaller than the true market, then flows across metro areas could still occur within the same true market. I address this concern by showing that job-to-job flows take place across many different distances, so small changes in the degree of aggregation cannot eliminate the presence of substantial cross-market flows.

Table 1.2 shows the share of cross-metro job-to-job hires over different distances and geographic boundaries. I use the geodesic distance between the center of each MSA. A sizable share of J2J flows are to relatively nearby destinations: 21% are between metros less than 50 miles apart, and 32% are between metros less than 100 miles apart. However, note that about 35% of job switches take place over a distance of 250 miles or greater, so long distance moves are relatively common. About 58% of cross-metro J2J hires are within state, which leaves a substantial share of flows that are across states. About 28% of flows are across the nine Census divisions, and 20% are across the four Census regions.⁵

I also estimate the number of cross-metro job switches that are within the same commuting zone. Commuting zones are a data-driven aggregation of counties into labor markets based on commuting flows between them, which makes them a prime candidate for an alternative labor market definition. I use the Census Bureau's delineation files for Core based statistical areas (CBSAs) and the USDA Economic Research Service's definitions for the 2000 vintage of commuting zones to make a county-level crosswalk between metro areas and commuting zones. There is no strict hierarchy between CBSAs and commuting zones: some metro areas contain or cross multiple commuting zones, and some commuting zones

⁵The four Census regions are Northeast, South, Midwest, and West. Each region is further divided into divisions.

Table 1.2: Share of Cross-Metro Job-to-Job Flows Within

Distance		Geography	
25 miles	10.87%	Commuting Zone:	
50 miles	20.98%	Lower Bound	6.43%
100 miles	32.75%	Upper Bound	11.59%
250 miles	65.01%	State	58.12%
500 miles	90.00%	Division	72.40%
1000 miles	99.13%	Region	80.11%

Notes: Share of average quarterly job-to-job flows across metro areas, 2003-2017. Lower bound for CZ includes flows between single-CZ MSAs. Upper bound includes flows between MSAs with any overlap in CZ. Source: J2JOD, U.S. Census Bureau

contain or cross multiple metro areas, so I cannot obtain an exact figure for the share of cross-metro flows that are within CZs. However, I can do a bounding exercise. For the lower bound, I count a cross-metro flow as within-CZ if the two metro areas contain only one commuting zone that is common between them. For an upper bound, I count any flow between metro areas with any overlap in commuting zones as within-CZ. At most, 11.5% of cross-metro flows are within the same commuting zone.

Lastly, I measure nonlocal labor demand using data on the rate of J2J outflows, which has substantial variation across MSAs. Figure 1.2 shows J2J outflows as a rate of employment and as a share of total J2J flows, plotted against total employment in each MSA. The heterogeneity is most evident when looking at J2J outflows as a share of total J2J flows, which range from around 20% at the low end to greater than 60% at the high end. Interestingly, there is a strong, negative relationship between J2J outflows and total MSA size. This is to be expected: the total employment of an MSA is endogenously determined by its productivity and local amenities, which contribute to local absorption of J2J flows. Scale effects in matching efficiency or the degree of assortative matching could further contribute to greater local absorption and higher wages in larger MSAs (Dauth et al., 2018; Card et al., 2021). In other words, places with a high outflow rate tend to be smaller MSAs, which economic geography models would predict are less productive and have lower average wages. This is one of any number of MSA characteristics that could contribute to spurious correlations between outward labor mobility and wage levels. In evaluating the effect of nonlocal labor demand on mobility and wages I use first differences, which helps to control for the

influence of MSA size and other unobserved MSA characteristics.

Figure 1.2: Cross-MSA Job-to-Job Flows: MSA-Level Variation

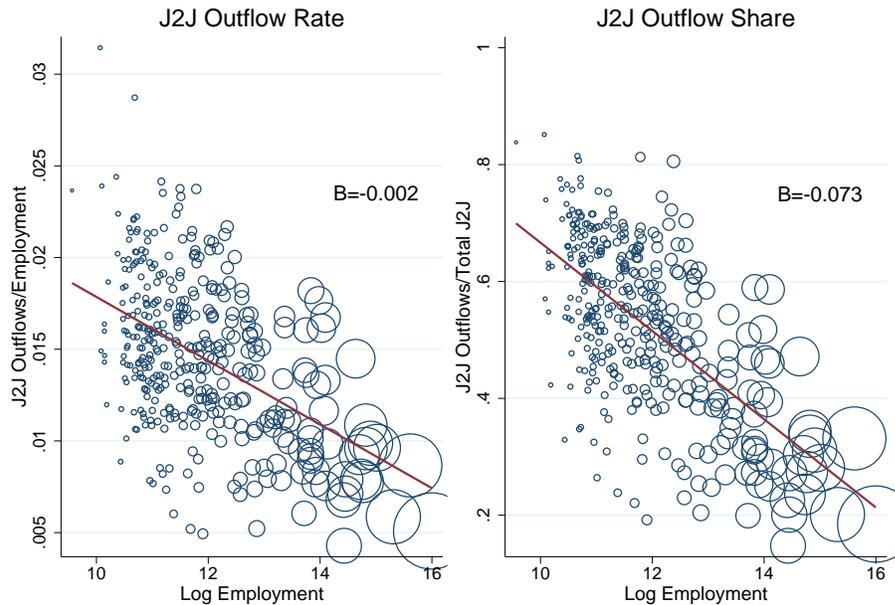


Figure 1.2: Average quarterly J2J outflow rate and J2J outflows as a share of total J2J flows, plotted against Log MSA-level employment. Marker size indicates total MSA-level employment. Average of quarterly values for the year 2015. Data: J2JOD, QWI

To summarize, Job-to-job flows across metro areas are a substantial share of all job-to-job flows. To the extent that job switching reflects workers' outside options, this suggests that non-local jobs are an important component of workers' labor market opportunities. How do changes in job opportunities in other labor markets through job switching and wage renegotiation? In the following section, I detail methodology for evaluating the contribution of non-local job opportunities on the geographic reallocation of labor and on wage growth through improvement in outside options.

1.4 Nonlocal Labor Demand, Labor Mobility, and Earnings Growth

In this section, I develop measures of changes in non-local outside options and present the empirical approach to use estimate their effect on 1) geographic labor mobility, and 2) earnings growth. The measures of non-local outside options are based on exposure

through historic job-to-job flow networks to year-over-year changes in hires in other MSAs. Intuitively, stronger hiring growth in other labor markets increases the likelihood of a local job searcher receiving a nonlocal outside offer, which the worker can exercise by switching jobs or leverage to renegotiate wages at their current job. I formalize this intuition by showing that the growth rate of hires is a proxy for the job finding rate, though an imperfect one as hiring also reflects changes in non-local labor supply. However, it is plausible to assume that labor demand shocks only influence hiring growth through the job finding rate, which motivates the instrumental variables strategy I use based on the Network IV specification of [Schubert \(2021\)](#). With this approach, shift-share instruments ([Bartik, 1991](#)) are used to identify exogenous variation in labor demand in all other MSAs aside from the local labor market. This instrument identifies non-local labor demand shocks, rather than labor supply shocks, and in particular it identifies shocks that are exogenous with respect to local labor demand.

The estimated response of worker mobility to non-local labor demand shocks validates my assertion that my measures of changes in non-local labor demand do indeed reflect changes in workers' outside options. They also illustrate demand-driven geographic reallocation of labor through job-to-job flows, which is an important element of aggregate productivity and wage growth. The estimated response of earnings growth, meanwhile, tests for the effect of non-local labor demand on wages for job-stayers through an outside options channel, and thus illustrates the potential for nonlocal labor market opportunities to reduce local employer market power.

1.4.1 Nonlocal Outside Options and the Job Finding Rate

In labor search models, workers' job opportunities and outside options are determined in part by the job finding rate, or the probability that a job searcher finds employment. When currently employed workers receive an outside offer, they can either accept the offer or leverage their improved outside option to renegotiate wages with their current employer. This source of intra-firm competition plays a major role in wage setting, and is likely more quantitatively important than the value of non-employment in determining wages ([Cahuc et al., 2006](#)).

In a multi-region environment, the job finding rate can be thought of as the sum of a number of bilateral job finding rates, each representing the probability that a worker in market m finds a job in market n . I want to evaluate workers' nonlocal outside options using precisely these bilateral job finding rates. However, job finding rates are not directly observable, so in this section I demonstrate that changes in hires and historic job-to-job flow rates can be used as proxy measures for bilateral job finding rates.

Let the local job-finding rate for market n in time t be denoted ϕ_{nt} . I assume that bilateral job-finding rate in market n for workers in market m is given by ϕ_{nt} discounted by a scalar factor α_{nm} . This factor can represent information frictions that reduce the effectiveness of searching for jobs in market n for a worker currently in market m , as well as moving costs, differences in industry or job composition, and local amenities or other factors that affect the willingness of workers in market m to accept jobs that they encounter in market n . The total nonlocal job finding rate, ϕ_m^{NL} , is the sum of $\alpha_{nm}\phi_{nt}$ over all nonlocal destination markets, and the percentage change in the nonlocal job finding rate can be approximated by

$$\Delta \ln \phi_m^{NL} = \sum_{n \neq m} \alpha_{nm} \ln \frac{\phi_{nt}}{\phi_{nt-1}}$$

Note that both α_{nm} and ϕ_{nt} are unobserved. However, hires and bilateral job-to-job flows can serve as proxies for the job finding rate and the bilateral discount factor. For the former, note that hires of local workers in market n is the effective population of local job searchers s_{nt} multiplied by the job finding rate.⁶ The log change in hires is then given by

$$\ln \frac{Hires_{nt}}{Hires_{nt-1}} = \ln \frac{\phi_{nt}}{\phi_{nt-1}} \ln \frac{s_{nt}}{s_{nt-1}} \quad (1.1)$$

In other words, the growth rate of hires reflects both changes in labor supply (s_{nt}) as well as changes in job finding rates (ϕ_{nt}). It is important to note that changes in labor supply can effect both ϕ_{nt} and s_{nt} , while the effect of labor demand is only reflected in ϕ_{nt} .

⁶Under the assumption that bilateral job finding rates differ from the local job finding rate by a scalar factor, local hires can be used here without loss of generality.

In addition, note that job-to-job flows from market m to n are given by the product of the effective searcher population s_{mt} and the bilateral job finding rate $\alpha_{nm}\phi_{nt}$. Under the assumption that the bilateral friction is not time-varying, J2J flows from a base period can be written as: $J2J_{nm}^* = s_m^* \alpha_{nm} \phi_n^*$. Combining these expressions and summing over destinations, nonlocal hires growth can be expressed as:

$$\sum_{n \neq m} J2J_{nm}^* \ln \frac{Hires_{nt}}{Hires_{nt-1}} = \sum_{n \neq m} s_m^* \alpha_{nm} \phi_n^* \ln \frac{\phi_{nt}}{\phi_{nt-1}} \ln \frac{s_{nt}}{s_{nt-1}} \quad (1.2)$$

This shows that the J2J-weighted sum of hiring growth across all nonlocal MSAs is a candidate proxy for the change in the nonlocal job finding rate, though it differs from an ideal measure in several ways. First, The bilateral terms used to weight hires growth contain terms reflecting initial labor supply and demand in the origin and destination, respectively. More importantly, growth in nonlocal hires reflects changes in nonlocal labor supply as well as changes in the nonlocal job finding rate. However, note that changes in labor demand affect hiring only through the job finding rate. This is a key motivation for the instrumental variables strategy I outline below: isolating variation in hiring to changes in labor demand identifies changes in the job finding rate.

1.4.2 Measures of Nonlocal Hires Growth

Following equation (1.2), I construct empirical measures of nonlocal hiring growth using data from the J2JOD and QWI. The data cover 2003-2019 at a quarterly frequency. For each destination, I take the year-over-year (4 quarter) difference in log hires. I use year-over-year differences, rather than quarterly differences, for two reasons. First J2J flows and hires are strongly seasonal, and the seasonal trends differ across variables and across MSAs, making it difficult to implement any seasonal adjustment procedure. Year-over-year growth rates are a more transparent way of seasonally adjusting these labor market indicators. Second, the data from the Current Population Survey only permits construction of annual wage growth measures, which I match to quarterly, year-over-year of nonlocal hiring growth.

To aggregate nonlocal hires growth, for each origin MSA I use the annual stable job-to-job

outflow rate as a percent of local employment from 2003 to weight nonlocal hires.⁷ Using historic J2J flows avoids the potential endogeneity of exposure through job-to-job flows to labor demand growth in other markets, and means that the time-variation in measured non-local demand is strictly due to changes originating in other MSAs. Using stable job-to-job flows (those between employment spells observed in at least 3 quarters) limits the effect of seasonal and other temporary employment on the estimated job-to-job flow connections between MSAs. In particular, note that seasonal employment cycles (e.g., working in city A in the winter and city B in the summer) do not contribute to stable J2J flows by construction, as it is not possible to hold multiple seasonal jobs that last for three quarters over a single year. Thus, my measure of nonlocal hiring growth from MSA m at time t is:

$$\Delta \text{NL Hires}_{mt} = \sum_{n \neq m} J2J_{mn}^{2003} \Delta^{t,t-4} \ln \text{Hires}_{nt}$$

I construct this measure at the MSA and the MSA-by-education level. For the latter, the growth rate of hires in the destination and the job-to-job outflow rates used for aggregation are constructed within bins of educational attainment (less than high school, high school, associate's degree or some college, and a 4-year college degree or greater). Note that education is not available or imputable for all observations in the underlying microdata. I omit observations with missing education when constructing MSA-by-education measures.

$$\Delta \text{NL Hires}_{mte} = \sum_{n \neq m} J2J_{mne}^{2003} \Delta^{t,t-4} \ln \text{Hires}_{nte}$$

⁷The sum of the weights is equal to the total J2J outflow rate as a percentage of employment in 2003 for MSA m . Arguably a more appropriate measure uses weights that are the share of total J2J outflows so the shares sum to one. The measure I construct, however, includes variation in the total J2J outflow rate, which varies across MSAs and influences exposure to any subsequent shocks; more remote places have fewer J2J outflows at all, and the nonlocal labor market is simply a weaker source of job opportunities. In this sense, shares that sum to one implicitly overweight more remote places. A related issue is the “incomplete shares” problem raised by [Borusyak et al. \(2020\)](#). In their analysis, they show that this source of variation can bias estimates as remoteness or other features that influence total outflows is likely correlated with long term outcome variables like wage growth and mobility rates directly. My estimates address this by including MSA fixed effects, which absorb time-invariant factors such as the total J2J outflow rate in 2003. Additionally, I find that including lagged measures of the total J2J outflow rate as a control has little quantitative effect on the estimates I present.

1.4.3 Empirical Framework

I use non-local hiring growth to evaluate the effect of changes in non-local outside options on two outcomes: job-to-job outflows and wage growth for job stayers. In general, I wish to estimate models of the following form:

$$\Delta \ln V_{mt} = \beta_0 + \beta_1 \Delta \text{NL Hires}_{mt} + \Xi X_{mt} + \delta_m + \kappa_t + u_{mt} \quad (1.3)$$

Where V_{mt} is an outcome variable for MSA m at time t (job-to-job outflows or wage growth), X_{mt} is a vector of MSA-level controls, and δ_m and κ_t are MSA and time fixed effects. The dependent variables and the measure of nonlocal hires are expressed in first differences, which cancels out any time-invariant MSA-level characteristics that influence both the level of exposure to nonlocal hires and the outcome variables. I additionally include MSA fixed effects, which help control for measurement error in the construction of nonlocal hiring growth as a proxy for changes in the nonlocal job finding rate. In equation (1.2), this measurement error is due to local labor supply s_m^* and the weighted sum of non-local job finding rates ϕ_n^* in the base year, which contribute to the J2J outflow weights used to aggregate growth in nonlocal hiring but are not representative of the underlying bilateral connection. Additionally, the MSA fixed effects control for the sum of the base-year J2J weights, addressing the “incomplete shares” bias (Borusyak et al., 2020), as well as any time-invariant MSA characteristics that influence long-run J2J outflows and wage growth, such as industry composition.

It is unlikely that OLS estimates of β_1 will produce causal effects. First, it is possible that growth in non-local hires reflects non-local labor supply shocks, rather than labor demand shocks, and second, it possible that these nonlocal labor demand shocks shocks not independent with respect to local labor demand shocks. If non-local hiring growth is driven by non-local labor supply (in either the destination MSA or through labor mobility from other MSAs), this would tend to bias estimates of mobility and earnings downwards. Intuitively, increases in labor supply in destination markets lower the job finding rate and wages, reducing rather than improving outside options. If, on the other hand, non-local hiring growth is due to increased labor supply from local outflows, the negative labor supply shock to

the local market would increase the local job finding rate, improving local outside options. This would bias the effect of non-local demand shocks on earnings and mobility upwards.

It is also possible that labor demand shocks for MSAs with strong J2J network connections are correlated. As job-to-job flows are influenced by geographic proximity and industry similarity, it is likely that local labor demand shocks are correlated within MSAs that experience greater labor flows between them. This has an ambiguous effect on the mobility estimates: on one hand, increased local labor demand would offset the effect of increased non-local demand on job-to-job outflows by absorbing job switches locally, but on the other hand, increased local labor demand could result in poaching of workers from nearby MSAs, resulting in increased vacancies and higher gross labor demand in destination MSAs. However, correlated labor demand shocks would unambiguously bias the estimated effect of non-local labor demand on wage growth upwards, since an improvement in local outside options leads to increased local wages. While in all specifications I control for various measures of local labor market conditions, the threat of correlated shocks conditional on these equilibrium outcomes remains.

In the following section, I outline an instrumental variables approach that allows me to identify the causal effects of nonlocal outside options on outward labor mobility and wage growth. In this specification, I use Bartik “shift-share” instruments to identify exogenous variation in non-local labor demand in all other cities. This ensures identification of non-local labor demand shocks, rather than nonlocal labor supply shocks, and in particular nonlocal labor demand shocks that are exogenous with respect to local labor demand.

1.4.4 Network Instruments

To produce causal estimates of the effect of nonlocal outside options on J2J outflows and wage growth, I use a Network IV specification based on [Schubert \(2021\)](#). This strategy is an extension of the [Bartik \(1991\)](#) “shift-share” approach that leverages cross-city labor mobility rates to construct an instrument that identifies exogenous variation in nonlocal labor demand.

In the [Bartik \(1991\)](#) approach, national industry growth rates are interacted with local in-

dustry shares to create a measure of local exposure to national industry-level shifts in labor demand. With the Network IV, the shift-share instruments are used to identify exogenous shifts in nonlocal labor demand. Mechanically, the weighted sum of nonlocal hiring growth is instrumented for using the weighted sum of shift-share instruments for hires in each nonlocal destination. As recent literature shows, shift-share instruments leverage differential exposure to a common shock, and the validity of the instrument relies on exogeneity of the initial shares (Borusyak et al., 2020; Goldsmith-Pinkham et al., 2020). For the Network IV, the shares that determine local exposure to national industry shocks are the J2J-share-weighted industry shares in nonlocal labor markets. I will discuss below the plausibility of the exogeneity of these shares given the construction of the instruments and regression specification.

Local Shift-Share Instruments

For each MSA, I first construct Bartik (1991) shift-share instruments using data from the Quarterly Workforce Indicators (QWI) on MSA- and national-level employment and hires for 3-digit NAICS industries. The Bartik IV for locality n in period t , B_{nt} , is the sum of local industry employment shares s_{in} times the year-over-year national industry hiring growth.

$$B_{nt} = \sum_{i \in I} s_{in}^{2003} \Delta^{t,t-4} \ln \text{Hires}_{it} \quad (1.4)$$

I fix the local employment share of industry i in MSA n s_{in} at the average value in 2003. Using fixed historic industry shares avoids potential endogeneity as local employment responds to national industry trends. I compute national industry growth on a leave-one-out basis: the local change in hires in industry i is subtracted from the national change in hires. This avoids a mechanical correlation between local and national hiring growth.

Network Instrument

Under certain conditions, these shift-share instruments plausibly identify exogenous shifts in MSA labor demand. The Network IV uses a weighted sum of these instruments to

identify local exposure to exogenous shifts in nonlocal labor demand:

$$NB_{mt} = \sum_{n \neq m} J2J_{mn}^{2003} B_{nt}^{-m} \quad (1.5)$$

The weights I use to aggregate the Network IV are identical to those used in constructing the measure of nonlocal hires, so the instrument relies on the identifying assumption that historic cross-MSA J2J flow rates are indicative of current exposure to changing nonlocal labor market conditions. Job-to-job flow rates are fixed at their level in 2003 to limit the influence of endogenous labor mobility in response to labor demand shocks, and I use stable job-to-job flows as a percentage of stable employment to limit the influence of flows driven by seasonal or temporary work.⁸ In addition, I adjust each destination's shift-share instrument to remove the contribution of the origin locality m to the national hiring growth in each industry. This leave-2-out construction ensures that hires growth in the origin does not mechanically influence the Network IV.

I also construct an education-specific Network IV. In this case, the MSA-level shift-share measures are still constructed over all workers, but I weight exposure using J2J outflow rates specific to education-level bins⁹.

$$NB_{mte} = \sum_{n \neq m} J2J_{mne}^{2003} B_{nt}^{-m} \quad (1.6)$$

Note that the typical shift-share IV employed in the literature uses growth in employment or wage bills, while in this application I use national hiring growth. Changes in employment

⁸The sum of the weights is the total job-to-job outflow rate in 2003 for MSA m . As [Borusyak et al. \(2020\)](#) note, shares that do not sum to one in shift-share instruments can lead to an identification problem: the sum of the weights can be correlated with the outcome variables. In this case, it is plausible that the total job-to-job outflow rate influences average wage growth. They recommend controlling for the sum of the weights for robustness to this form of endogeneity. In the specification I use, the sum of the total weights is captured by the inclusion of MSA fixed effects. As a robustness check, I include a two-year lag of the total J2J outflow rate as an additional control, and find this makes little quantitative difference to the main results.

⁹I find that using overall, rather than education specific, local industry employment shares and changes in national industry hires, results in greater predictive power in the first stage for education specific nonlocal hires by education. This result is specific to using 3-digit NAICS industries, where local education-by-industry employment shares have more numerous empty or nondisclosed cells.

and wages are measures of net labor demand, and include variation due to separations as well as hires. In this case, since I am instrumenting for changes in gross labor demand (hires growth) national hiring growth is a more natural candidate for an IV. Indeed, in first-stage regressions I typically find that the hiring Network IV is a stronger predictor of local hires than similarly constructed employment or earnings Network IVs. That said, the employment Network IV has additional predictive power for local hires conditional on the hiring IV, so I explore specifications that include both instruments.

Identification

Shift-share instruments have been widely used since their introduction in [Bartik \(1991\)](#), and recent work has formally demonstrated the conditions that permit valid causal inference ([Borusyak et al., 2020](#); [Goldsmith-Pinkham et al., 2020](#)). The industry shares for each MSA measure differential exposure to common shocks (national industry growth rates), and the exclusion restriction requires exogeneity of the initial shares with respect to unobserved shocks to the dependent variables.

The Network IV relies on a similar justification. As [Schubert \(2021\)](#) shows, the exclusion restriction for the Network Bartik IV is

Network Exclusion Restriction: Assuming the network instrument is relevant, then the IV estimate of the effect of nonlocal labor demand on is consistent if and only if

$$\sum_{t \in T} \sum_{i \in I} s_i g_{it} \bar{\eta}_{it}^{V, \perp} \xrightarrow{P} 0$$

Where s_i refers to the national hiring share of industry i , g_{it} refers to the time-varying national hiring growth rate of industry i , and $\bar{\eta}_{it}^{V, \perp}$ is variable V (job-to-job outflow growth or earnings growth) after being residualized over all control variables and averaged over all cities, weighted by their J2J network exposure to national industry shocks.

The exclusion restriction is satisfied if unobserved shocks to the local outcome variables (labor mobility and wage growth) are not systematically correlated with local exposure through J2J networks to industries that have systematically higher national industry growth

trends. In other words, the initial shares that need to be exogenous (with respect to local labor demand shocks) are the J2J-outflow-weighted sum of nonlocal industry employment shares.

While nonlocal industry shares are indeed more plausibly exogenous than local industry shares, the job-to-job flow weights are an equilibrium outcome, so concerns of endogeneity remain. It is likely that MSA-level characteristics such as industry composition, geographic centrality, and market size are correlated with job-to-job outflows and wages. However, as I estimate models where the outcome variables are *differences* in local J2J outflows and wages, the relationship between endogenous J2J outflows and the *levels* of these outcome variables are not a concern.

Two features of my model specification further mitigate concerns of endogeneity. First, I include MSA fixed effects, which control for the effects of initial industry composition and other latent variables such as geographic centrality on the long-run average of growth rates of local J2J outflows and wage growth. Lastly, in all IV specifications I apply the “double-Bartik” method by including the local Bartik labor demand measure as a control variable (Schubert, 2021; Chodorow-Reich and Wieland, 2020). This controls for time-varying local exposure to national industry trends, so identification of the effects of nonlocal demand is thus based on the the difference between local and J2J-weighted nonlocal industry composition.

1.4.5 Estimating Equations

Nonlocal Demand and J2J Outflows

To test for the effect of nonlocal labor demand on local job-to-job outflows I estimate the following model using data from the J2JOD and QWI.

$$\Delta \ln \text{J2J Outflow Rate}_{mt} = \beta_0 + \beta_1 \Delta \text{NL Hires}_{mt} + \beta_2 B_{mt} + \Xi X_{mt} + \delta_m + \kappa_t + u_{mt} \quad (1.7)$$

Where $\Delta \text{NL Hires}_{mt}$ is instrumented using NB_{mt} . The Network IV uses base year weights

from 2003, and nonlocal hires growth uses J2J shares that are lagged one year; this rescales the estimates to be in terms of current exposure to nonlocal demand while ensuring that the identifying variation is based on historic exposure. The dependent variable is the year-over-year log difference in the quarterly rate of stable job-to-job outflows as a percentage of local employment. The coefficient of interest is β_1 : the effect of growth in non-local labor demand on job-to-job outflows. As the growth rate of job-to-job outflows and nonlocal hires are expressed in logs, β_1 can be interpreted as an elasticity. B_{mt} is MSA m 's local shift-share measure, which controls for time-varying local exposure to the same national industry shocks captured in NB_{mt} . The vector X_{mt} includes MSA-level controls intended to capture changes in local labor market conditions, including year-over-year growth rates of local hires, employment, and total payrolls. These local indicators control for local labor supply and demand factors that may influence the outcome variables. I additionally explore the possibility that these controls absorb the effect of nonlocal demand and thus affect β_1 by reporting estimates with local controls excluded. The estimation sample period is 2005-2019, which avoids contact with the base year rates in 2003 used for construction of non-local labor demand. Table A1.1 reports summary statistics for the MSA-level estimation sample.

In my preferred specification, I additionally control for 4 quarterly lags of these variables, as well as 4 lags of the local Bartik measure B_{mt} and the Network IV NB_{mt} . The motivation for controlling for lags of the Network IV is twofold. First, the literature on shift-share instruments highlights persistence in the both the weights used to measure exposure to the common shocks and the response to the shocks, which can lead to contamination of estimates by the responses to past shocks (Jaeger et al., 2018; Goldsmith-Pinkham et al., 2020). In our case, it is possible that shocks to nonlocal labor demand in the past induce persistent responses of job-to-job outflows, and are correlated with current shocks to nonlocal labor demand.¹⁰ Jaeger et al. (2018) and Caballero et al. (2021) suggest controlling for lagged values of the shift-share instrument to address persistent responses to prior shocks. This concern also motivates including lags of the local controls and local Bartik measure: it is possible that persistent responses to local labor supply and demand shocks bias estimates

¹⁰In some specifications, omitting the controls for lags of the excluded instruments leads the model to fail a key falsification test: I find an estimated effect of current demand shocks on *lagged* outcomes. When controlling for lagged values of the IV, the model passes this falsification test.

of the effect of current shocks. Second, recall that the J2J outflow data are quarterly, but the variables are differenced on a year-over-year basis to address seasonality. Controlling for lagged values of the IV means that the identifying variation is the unexpected quarterly innovations in the IV, so as a matter of interpretation is closer a quarterly nonlocal demand shock. This allows me to more precisely evaluate the dynamics of the response of J2J outflows with respect to nonlocal demand.

Nonlocal Demand and Wage Growth

I also estimate the effect of nonlocal labor demand on wages. Recall that the structure of the CPS Earner Study means that I have two earnings observations per worker that are 12 months apart, which I use to construct annual wage growth. For details on the construction of wage growth using the CPS, see Section 3. I restrict the sample to workers that report no recent job switches, which allows me to identify the effect of nonlocal outside options on wages independent of firm and worker effects.

In my preferred specification, I estimate models using education-specific measures of non-local hires and the Network IV. In this case, the estimating equation is

$$\Delta \ln \text{Wage}_{it} = \beta_0 + \beta_1 \Delta \text{NL Hires}_{mte} + \beta_2 B_{mt} + \Xi X_{mt} + \eta Z_{it} + \delta_{me} + \kappa_t + u_{it} \quad (1.8)$$

Where $\Delta \text{NL Hires}_{mte}$ is instrumented for using NB_{mte} . The Network IV uses base year weights from 2003, and nonlocal hiring growth uses J2J shares that are lagged one year; this rescales the estimates to be in terms of current exposure to nonlocal demand while ensuring that the identifying variation is based on historic exposure. The dependent variable is year-over-year growth in hourly wages for person i working in MSA m , X_{it} is a vector of individual controls including age, race and ethnicity, education level, and 2-digit occupation and industry fixed effects, and Z_{mt} is a vector of time-varying MSA-level controls, including year-over-year growth in local hires, employment, and payrolls, and the local unemployment rate. These local indicators control for local labor supply and demand factors that may influence the outcome variables. I additionally explore the possibility that these controls absorb the effect of nonlocal demand and thus affect β_1 by reporting estimates

with local controls excluded. δ_{me} are fixed effects for MSA-by-education-level bins, so coefficients can be interpreted as the within MSA effects of changes in non-local demand. For the education-specific model I cluster standard errors at the MSA-by-education-bin level. The timing of the CPS data is monthly, so I match monthly worker-level observations of wage growth to quarterly measures of nonlocal hires, the Network IV, and other MSA-level controls. The estimation sample period is 2005-2019, which avoids contact with the base year rates in 2003 used for construction of non-local labor demand. Table A1.2 reports summary statistics for the CPS sample used in estimation, and Table A1.3 reports the distribution of individual observation counts at the MSA and MSA-by-education level ¹¹.

One issue with using hires to construct measures of nonlocal labor demand or the Network IV is that hires are a flow measure and thus are not cumulative. It is possible that the year-over-year change in hires does not fully capture the cumulative change in labor demand over the previous year. For the wage growth estimates, this is particularly relevant, as I only observe cumulative growth in wages over the previous year, and it is possible that current wages reflect labor demand shocks that occurred prior quarters. To address this, I estimate alternative versions of equation 1.8 using measures of nonlocal hires, the local shift-share variable, and the Network IV that use cumulative YoY change in hires over the previous year - essentially a rolling annual growth rate of hires. In addition, I explore including a second Network IV constructed with changes in employment as another method of capturing the cumulative variation in nonlocal demand.

1.5 Results

1.5.1 Nonlocal Demand and J2J Outflows

In this section I detail estimates of the effect of nonlocal labor demand shocks on job-to-job outflows. These estimates serve two functions. First, the outward mobility response of workers serves as validation that the measures of changes in non-local labor demand reflect

¹¹Sample counts in the CPS can be low for smaller MSAs, particularly when the variation used to estimate the main coefficients is at the MSA-by-education level. This raises concerns about the precision of estimates. As a robustness check, I re-estimate the main specification restricting the sample to MSAs or MSA-by-Education bins with observation counts that are above the 25th or 50th percentile. Table A1.6 reports results. I find excluding low-observation MSAs makes little difference to the main results, and indeed slightly increases the estimated effect of of nonlocal labor demand on wage growth

changes in workers' outside options. Second, these estimates are informative about the extent of demand-driven geographic reallocation of labor through job-to-job flows, which is a source of productivity and wage growth as well as an important mechanism in regional adjustment (Blanchard and Katz, 1992).

Table 1.3: Nonlocal Demand and J2J Outflows

	Dependent Variable: $\Delta \text{Log J2J Outflow Rate, YoY}$		
	2SLS (1)	IV: Hires (2)	IV: Hires & Employment (3)
$\Delta \text{ Nonlocal Hires}$	0.369*** (0.0493)	0.306*** (0.0956)	0.254*** (0.0755)
Local Bartik: Hires	-0.0699 (0.0598)	-0.0682 (0.0600)	-0.0738 (0.0601)
Local Bartik: Employment			-0.246*** (0.0748)
$\Delta \text{ Log Employment}$	-1.296*** (0.104)	-1.292*** (0.105)	-1.279*** (0.104)
$\Delta \text{ Log Hires}$	0.0598*** (0.0196)	0.0740*** (0.0234)	0.0861*** (0.0206)
$\Delta \text{ Log Payroll}$	0.0144 (0.0259)	0.0174 (0.0263)	0.0204 (0.0261)
N	18415	18415	18415
First-stage F-stat	1214.7	641.8	397.1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is quarterly data from 2005-2019. 2SLS estimates use base-year (2003) exposure to nonlocal hires to instrument for 1-year lagged exposure to nonlocal demand. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand. All models include MSA and date fixed effects and control for current and 4 quarterly lags of local changes in employment, hires, payroll, the local Bartik shift-share measure, and the excluded instrument(s).

Table 1.3 reports the main results from estimating equation (1.7). Column 1 contains OLS estimates, for which I find a coefficient of about .37. As nonlocal hires and growth in the J2J outflow rate are both in logs, this coefficient can be interpreted as an elasticity: a 1% increase in nonlocal hires is associated with a .37% increase in J2J outflows. While this result demonstrates a strong correlation between nonlocal hires and job-to-job outflows, OLS is unlikely to return causal estimates of the effect of nonlocal hires due to the mixing of labor supply and demand shocks and the potential for correlated local and non-local labor demand shocks (though in the case of labor mobility the direction of the bias is

ambiguous). Column 2 reports the IV estimates as detailed in the previous section. The estimated elasticity using Network Bartik instrument is .306, somewhat smaller than the OLS elasticity. This effect is statistically significant, and the strong F-statistic in the first stage (600) indicates that weak instruments are not an issue. This serves as the benchmark estimate of the elasticity of job-to-job outflows with respect to nonlocal labor demand.

To evaluate the economic significance of this estimate, consider the effects of a typical shock to nonlocal hires. After controlling for 4 quarterly lags of growth in nonlocal hires, the within-MSA standard deviation of current nonlocal hiring growth is about .057 log points (about two thirds the value of the within-MSA standard deviation of local hires). With an elasticity of .306, this means that a typical positive nonlocal demand shock leads to an increase in the J2J outflow rate of about 1.72%. To put this estimate into context, in column (1) of Table 1.4, I report the effect of local labor demand on local job switching estimated using the local shift-share measure as an instrument for YoY growth in local hiring. I find that the elasticity of local job switches with respect to local hires is about .59, or about twice the elasticity of cross-MSA job switches with respect to nonlocal labor demand. A typical shock to local hires growth (.088 log points) leads to a 5.5% increase in local job switches, implying that the response of job switches to nonlocal labor demand is about 30% of the effect size of a typical shock to local demand. In absolute terms, evaluated at the average quarterly J2J outflow rate of 1.41%, a 1.72% increase in J2J outflows translates to a .025 percentage points in the rate. So while the magnitude of the estimated elasticity is moderately sized, the contribution of nonlocal demand shocks in explaining geographic labor labor mobility is relatively small. Indeed, these estimates suggest nonlocal labor demand shocks explain no more than 10% of the variation in overall job-to-job outflows. This indicates that supply-side factors, individual migration decisions, or idiosyncratic variation in workers' labor market opportunities that are not associated with aggregate nonlocal labor demand may be larger drivers of J2J outflows.

In Column 3, I modify the benchmark specification by including a network IV constructed with employment growth rates as an additional instrument for nonlocal hires growth, as well as controls for a local employment shift-share variable and four quarterly lags of the employment network IV. While there are intuitive reasons to think that national industry

shocks to hires are a stronger predictor of nonlocal hires, national industry shocks to employment capture a different source of variation that may be informative nonetheless. The main results are robust to including an employment IV: the estimated elasticity is somewhat lower at .254, though I am unable to reject the null hypothesis of equality between this coefficient and the benchmark specification. Note that the lower first F-statistic of the first stage estimate when including the employment IV indicates that the hires IV alone is a stronger predictor, though the second stage coefficient is somewhat more precisely estimated when including the employment IV ¹².

Table 1.4: Labor Demand and Local Job-to-Job Flows

	Dependent Variable: $\Delta \text{Log Local J2J Rate}$				
	(1)	(2)	(3)	(4)	(5)
	IV: Local Bartik	2SLS	IV	2SLS	IV
$\Delta \text{ Local Hires}$.5932*** (.0749)				
$\Delta \text{ Nonlocal Hires}$		0.333*** (0.0286)	0.0246 (0.0397)	0.0409 (0.0298)	-0.0219 (0.0393)
Local Bartik			0.737*** (0.0689)		0.0841 (0.0661)
$\Delta \text{ Log Employment}$				-0.856*** (0.0635)	-0.841*** (0.0639)
$\Delta \text{ Log Hires}$				0.594*** (0.0224)	0.602*** (0.0230)
$\Delta \text{ Log Payroll}$				0.0609 (0.0427)	0.0643 (0.0429)
N	19847	19847	19847	19847	19847
First-stage F-stat	226.86	4439.9	1169.3	3506.1	1325.2

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is quarterly data from 2005-2019. Column 1 reports estimates for the effect of local labor demand on local job switching using a local shift-share IV. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand. 2SLS estimates use base-year (2003) exposure to nonlocal hires to instrument for 1-year lagged exposure to nonlocal demand. All models include MSA and date fixed effects and four quarterly lags of all controls and excluded instruments.

Spatial correlation in MSA-level labor demand raises concerns for causal interpretation of these estimates. As a falsification test, I estimate the effect of nonlocal labor demand

¹²Table A1.4 reports the first-stage of IV estimates for models using hires and/or employment network instruments. The coefficients for the association between nonlocal hiring growth and the hires or employment IV are quantitatively similar.

shocks on local job switching: if shocks are spatially correlated, then nonlocal labor demand should be associated with increases in local job switching behavior due to the correlation with local labor demand. Table 1.4 reports the results of this exercise. In Column (2), I report OLS estimates omitting controls for local labor market conditions. I find an elasticity of local J2J with respect to nonlocal hires of about .33, indicating that spatial correlation is indeed a concern. In Column (3), I report IV estimates of the same elasticity that control for the local shift-share measure. The network IV is effective at limiting the role of spatial correlation, as the estimated elasticity is near zero and not statistically significant. When controlling for local labor market conditions, the OLS elasticity is similarly small and insignificant, and the IV elasticity is actually slightly negative.

Dynamic Response of J2J Outflows

To illustrate the dynamic response of J2J outflows with respect to nonlocal labor demand shocks I estimate impulse response functions using a local projection method. To produce the local projection estimates, I simply regress the k quarter forward value of the dependent variable on the current values of the independent variables. Table 1.5 reports results up to 3 quarters out. The top panel reports the YoY increase in J2J outflows in quarter $t + k$, and the lower panel reports the cumulative sum of YoY growth in J2J outflows from quarter t to quarter $t + k$. In the initial period, the contemporaneous increase of J2J outflow with respect to nonlocal labor demand is .305.¹³ In the following quarter, J2J outflows remain elevated on a year-over-year basis, though the response has declined to .217. Two quarters after the initial shock, the increase is smaller still and not statistically different from zero. By the third quarter, J2J outflows actually decrease, which suggests that nonlocal labor demand shocks pull forward some J2J outflows that would have occurred anyway. The cumulative response of J2J outflows levels up to 3 quarters following the initial shock is about .56, so the long run response of geographic labor mobility is substantially higher than the increase observed in the initial quarter. To put this estimate into context, a typical shock to nonlocal hires (.057 log points) results in a 1.76% increase in J2J outflows contemporaneously, and a 2.98% cumulative increase in J2J outflows over the 3 quarters after the initial shock.

¹³Note the slight numerical difference between this coefficient and the benchmark estimate. The local projection sample ends in 2017q2 rather 2018q4 to preserve comparability of coefficients of across different forward time horizons

Table 1.5: Dynamic Response of J2J Outflows to Nonlocal Demand: Network IV Estimates

Dependent Variable: $\Delta \text{Log J2J Outflow Rate, } t+k$				
k=	0	1	2	3
$\Delta \text{ Nonlocal Hires}$	0.305*** (0.0943)	0.217*** (0.0682)	0.0791 (0.0680)	-0.0343 (0.145)
N	16282	16282	16282	16279
First-stage F-stat	647.3	648.0	648.0	645.9

Cumulative Response				
Dependent Variable: Cumulative $\Delta \text{Log J2J Outflow Rate, } t+k$				
k=	0	1	2	3
$\Delta \text{ Nonlocal Hires}$	0.305*** (0.0943)	0.523*** (0.130)	0.603*** (0.167)	0.569** (0.245)
N	16282	16279	16276	16270
First-stage F-stat	647.3	647.4	647.4	645.3

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is quarterly data from 2005-2019. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand. Impulse responses are estimated with local projection over quarterly future values. All models include MSA and date fixed effects and control for current and 4 quarterly lags of local changes in employment, hires, payroll, the local Bartik shift-share measure, and the excluded instrument. Standard errors clustered at the MSA level.

Heterogeneous Effects

Top Destinations

In this section, I consider heterogeneity in the impact of nonlocal labor demand shocks on job-to-job outflows. One plausible dimension of heterogeneity is in the degree of geographic connectivity between the origin and destination MSA. I analyze this in two ways. First, I distinguish J2J flows between MSAs that have overlapping commuting zones versus those that are strictly across commuting zones. Analyzing flows that are strictly across commuting zones serves as a useful robustness check: commuting zones are a data-driven aggregation of counties meant to approximate local labor markets, so if the main effects on outflows are driven by within-commuting-zone J2J flows, this would suggest that my estimates reflect the effects of local labor demand. In addition, the degree of heterogeneity

between the response to within- or across-commuting zone shocks provides some indirect information about the responsiveness of labor mobility through different channels: commuting to alternative destinations incurs lower fixed costs than migration, so we should expect labor demand shocks from destinations within commuting distance to have a greater effect on labor mobility. Second, I analyze the effect of labor demand shocks from the top historic destinations for each MSA. Intuitively, destinations with greater historic J2J flows are likely those where workers have more information about labor market conditions and incur lower moving or commuting costs if they were to accept a job, both of which could lead to an increased mobility response to changes in outside options. Furthermore, the distribution of J2J outflows is quite skewed: flows to top-5 destinations account for about 60% of an MSA's J2J outflows on average, and about 70% of the variation in nonlocal hires growth, so heterogeneous responses to shocks from these destinations is important for understanding the aggregate implications of nonlocal demand shocks.

To estimate the heterogeneous responses to demand shocks coming from different subsets of destinations, I estimate models where J2J outflows, nonlocal hires, and the Network IV are constructed using data only for each subset. For regressions focusing on a particular subset, I add controls for the Network IV for the excluded set of destination MSAs ¹⁴

Table 1.6 reports the results of these exercises. The top panel distinguishes outflows to other MSAs in the same CZ from outflows to MSAs in different CZs. Note that not all MSAs have overlapping commuting zones with another, and those that do are more centrally located, so I report results separately for those with and without overlapping commuting zones. For MSAs that do have CZ overlap, the elasticity of J2J outflows for within-commuting zone demand shocks is .405, compared to an elasticity of .321 for cross-commuting zone demand shocks. For MSAs with no other MSAs in their CZ, the estimated elasticity is .307. So while J2J outflows are about 30% more responsive to shocks from MSAs within the same commuting zone, the benchmark estimate is robust to only including labor demand that is strictly across commuting zone. The lower panel reports estimated

¹⁴One potential concern for the validity of this exercise is that labor demand shocks are more likely to be correlated for MSAs with the greatest historic bilateral J2J flows. As a check, Table A1.5 reports the estimated effects of nonlocal labor demand from MSA-specific top-5 destinations and non-top destinations on local job switching. In both OLS and IV models that control for local labor market conditions, the estimated elasticities are near zero and not statistically significant for both top and non-top destinations.

Table 1.6: Nonlocal Demand and J2J Outflows: IV Estimates, Closely Connected Destinations

Dependent Variable: Δ Log J2J Outflow Rate for Specified Destinations		
Destinations within or Across Commuting Zones:		
	MSAs w/o CZ Overlap	MSAs w/ CZ Overlap
Δ Nonlocal Hires, Within CZ	N/A	0.405*** (0.108)
Δ Nonlocal Hires, Across CZ	0.307** (0.121)	0.321** (0.140)
N	18355	18355
Top-5 and Non-Top-5 Destinations		
Δ Nonlocal Hires, Top 5	0.338*** (0.0812)	
Δ Nonlocal Hires, non-Top 5	0.186 (0.235)	
N	18415	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is quarterly data from 2005-2019. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand. All models include MSA and date fixed effects and control for current and 4 quarterly lags of local changes in employment, hires, payroll, the local Bartik shift-share measure, and the excluded instruments. Each row is a separate model that excludes the network IV for the denoted set of destinations, and controls for the complementary IV. J2J outflows are counted as within CZ if the destination MSA has any overlapping commuting zones with the origin. Top-5 destinations are the MSA-specific top 5 in the base year (2003).

elasticities for shocks originating in top-5 and non-top-5 destinations. The elasticity for top destinations is .338, somewhat greater than the benchmark elasticity. For other destinations, the elasticity is substantially lower at .186, and not statistically significant. These results suggest that the effect of nonlocal outside options on labor mobility is largely due to demand from clusters of MSAs with with strong labor market connections to the origin. Indeed, I find that the response of job-to-job flows to labor demand in top-5 destinations accounts for about 74% of the total labor mobility response.

Effect by Education Level

A number of papers including [Bartik \(2017\)](#), [Molloy and Smith \(2019\)](#), and [Notowidigdo \(2020\)](#) have noted that geographic mobility rates of less-educated workers are relatively

Table 1.7: Nonlocal Demand and J2J Outflows: Effect by Education Level

	Dependent Variable: Δ Log J2J Outflow Rate, YoY			
	2SLS	IV	2SLS	IV
	(1)	(2)	(3)	(4)
Δ Nonlocal Hires by Educ	0.382*** (0.0344)	0.292*** (0.0664)		
$\times < \text{HS}$			0.318*** (0.0380)	0.244*** (0.0627)
$\times \text{HS}$			0.438*** (0.0464)	0.344*** (0.0729)
$\times \text{Assoc \& Some Coll}$			0.431*** (0.0473)	0.323*** (0.0773)
$\times \geq \text{College}$			0.414*** (0.0553)	0.299*** (0.0916)
Local Bartik: Hires		-0.128*** (0.0414)		-0.130*** (0.0418)
N	73660	73660	73660	73660
First-stage F-stat	3748.2	1882.8	872.5	558.7

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is quarterly data from 2005-2019. 2SLS estimates use base-year (2003) exposure to nonlocal hires to instrument for 1-year lagged exposure to nonlocal demand. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand. All models include MSA and date fixed effects and control for current and 4 quarterly lags of local changes in employment, hires, payroll, the local Bartik shift-share measure, and the excluded instruments. Standard errors clustered at the MSA-by-education level.

low, both on average and in response to local shocks. To see if this holds for nonlocal shocks, I estimate the effect of education-specific shocks to nonlocal demand on J2J outflow rates by education level. One important caveat is that in J2J statistics and underlying LEHD data, education is mostly imputed. This changes the interpretation of heterogeneity by education in two ways. First, any differences in these estimates across groups will be shrunk due to attenuation bias. Second, the education estimates will pick up heterogeneity based on the worker characteristics used in the imputation procedure. In the LEHD, the imputation of education levels is based to a large degree on workers' observed earnings, so the effects that I report can also be interpreted as differences in the mobility response to nonlocal labor demand between high and low income workers.

Table 1.7 reports results. I find that for workers without a high school degree, the effect of

nonlocal demand shocks on J2J outflows is lower than for other education groups, with an estimated elasticity of .244. This is 10 percentage points less than the elasticity for workers with a high school degree, and the difference is statistically significant ($p=.0196$). For all other groups of workers (High school, Some College & Associates degree, Bachelors degree or higher), I find that the effects of nonlocal demand shocks on J2J outflows are quantitatively similar. This somewhat muted degree of heterogeneity by education level may be explained by the fact that education is mostly imputed in the underlying LEHD data, so the differences in these estimates across groups are shrunk due to attenuation bias.

1.5.2 Nonlocal Labor Demand and Wage Growth

In this section, I use the measures of non-local labor demand to estimate the effect of changes in outside options on earnings. Recall that the structure of the CPS Earner Study means that I have two earnings observations per worker that are 12 months apart. I restrict the sample to workers that report no recent job switches in order to isolate the effect of nonlocal outside options on wage growth through a renegotiation channel. As the dependent variable is wage growth over the previous year, which is a cumulative measure, my preferred specification uses measures of nonlocal hires and the Network IV that use cumulative growth rates. These measures are less responsive to quarter-to-quarter variation in nonlocal demand, but are appropriate to use for the wage growth estimates as I cannot determine the exact timing of wage increases.

To illustrate the difference in timing, Table 1.8 reports estimates of an annualized version of the labor mobility regressions. To annualize the data, I convert any flow measure (including J2J outflows and hiring growth) to be in cumulative year-over-year terms, and restrict the sample to the 4th quarter of the calendar year. By construction, the year-over-year change in any stock variable is already in cumulative terms. In general, the estimated elasticities are substantially lower in annual terms at around .15, about half the benchmark elasticity. This indicates that job switching behavior is responsive to labor demand shocks at a relatively high frequency, which is supported by the high initial response of J2J outflows to nonlocal demand in the impulse responses in Table 1.5. This evidence is consistent with [Caldwell and Harmon \(2019\)](#), who find using precisely identified worker-level measures of outside options that increased job switching is only observed in the initial month, with a negative

Table 1.8: Nonlocal Demand and J2J Outflows: Annualized Estimates

	Dependent Variable: Δ Log J2J Outflow Rate, Annual			
	2SLS (1)	IV: Hires (2)	IV: Hires (3)	IV: Hires & Emp (4)
Δ Nonlocal Hires Growth	0.247*** (0.0368)	0.145*** (0.0500)	0.149*** (0.0523)	0.147*** (0.0500)
Local Bartik: Hires	-0.0391 (0.106)	-0.0669 (0.107)	-0.0660 (0.107)	-0.0684 (0.111)
Lagged Network IV: Hires			-0.0186 (0.0578)	
Local Bartik: Employment				-0.0264 (0.226)
N	4600	4600	4600	4595
First-stage F-stat	623.6	313.6	306.9	183.5

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is annual data from 2005-2019. 2SLS estimates use base-year (2003) exposure to nonlocal hires to instrument for 1-year lagged exposure to nonlocal demand. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand. All models include MSA and year fixed effects and control for local annual change in employment, hires, payroll.

or null response in subsequent months. The shocks that I study are at a more aggregate level so a greater degree of persistence is to be expected, but higher frequency data on mobility or wages may reveal greater responsiveness still. Note, however, that while the annual elasticity is roughly half the size as the quarterly, annual shocks to nonlocal hires are larger, so the effect of a typical shock is similarly sized.

Table 1.9 reports benchmark results for the effect of nonlocal demand on wage growth. I report results for education-specific measures of nonlocal labor demand. The OLS estimate of the elasticity of wage growth with respect to nonlocal demand is .12. This effect of nonlocal demand on wage growth is robust to IV estimation with an elasticity of .11. In column 3, I accumulate nonlocal hires and the Network IV over the previous year, which more accurately reflects the cumulative nature of the annual wage growth measures. In this specification, the elasticity increases to .156, though it is less precisely estimated and not statistically different from the previous specification. Lastly, I include a Network IV constructed using changes in employment as an additional instrument, which results in an estimated elasticity of .146. Because the benchmark sample may include some local

job switches, I estimate models with the sample limited to respondents that report no job switches in any month in the second survey wave and that report working in the same industry and occupation in both wage observations. Table A1.11 reports results. I find that using this more precise criteria for identifying job stayers reduces the estimated elasticity to .106. This combined with the increase in standard errors causes the estimate to lose statistical significance, though quantitatively the main effect is similar to the benchmark specification.

Table 1.9: Nonlocal Demand and Earnings Growth

	Dependent Variable: YoY Δ Log Wage			
	2SLS (1)	IV: Hires (2)	IV: Cumul. Hires (3)	IV: Cumul. Hires & Emp (4)
Δ Nonlocal Hires by Educ	0.120* (0.0554)	0.112* (0.0662)	0.156* (0.0832)	0.146* (0.0794)
Local Bartik: Hires		0.420*** (0.142)		
Local Bartik: Cumul. Hires			0.0902* (0.0491)	0.432*** (0.138)
Local Bartik: Employment				-0.0825 (0.141)
Lagged IV: Hires		-0.0515 (.0682)		
Lagged IV: Cumul. Hires			-0.0281 (0.0825)	-0.0387 (0.0843)
Lagged IV: Emp				-0.0636 (0.267)
N	186179	186075	186075	186075
First-stage F-stat	14100.8	13417.0	4696.9	3502.8

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is 2005-2019. Standard errors clustered at the MSA-by-education level. All specifications restrict to same-MSA and same-employer observations, and observations with absolute changes in hours less than or equal to 10. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects.

To put these estimates into context, it is useful to compare them to the effect of local labor demand on wages. In Table 1.10, I report results of a regression of wage growth on local hiring growth, instrumented for using the local Bartiv IV. I find an elasticity of

.387. This is consistent with [Bartik \(2015\)](#), who finds real wage elasticities ranging from .25 – .45, depending on the initial level of unemployment and employment growth. The wage elasticity with respect to local demand is about three times the benchmark elasticity of wage growth with respect to nonlocal demand. This indicates that outside options and wages are more influenced by local labor market opportunities, as expected. Indeed, a typical shock to local labor demand (.131 log points) results in a 5.2% increase in wages, while a typical nonlocal shock (.110 log points) results in a 1.24% – 1.73% increase in wages.

Table 1.10: Local Demand Shocks and Wage Growth: Bartik IV

Dependent Variable: YoY Log Wage	
IV	
YoY Δ Local Hires	0.387** (0.175)
N	194684
First-stage F-stat	523.6

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is 2005-2019. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand. All specifications restrict to same-MSA and same-employer observations. Weekly earnings growth sample restricts to observations with zero change in hours. Wage growth sample restricts to observations with absolute changes in hours less than or equal to 10. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects.

The estimated elasticity for nonlocal demand shocks of .112 – .156 is large relative to some previous estimates of the effect of changes in outside options on individual wages for job stayers. [Caldwell and Harmon \(2019\)](#) find a quasi-elasticity of roughly .02, using variation in hiring at the firms of former coworkers, and [Lachowska et al. \(2021\)](#) finds a similarly sized elasticity using variation in wages offered at the secondary job of dual job holders. In contrast, my estimates are lower than previous estimates leveraging variance at the level of the local labor market. This includes [Caldwell and Danieli \(2018\)](#), who find a wage elasticity between .17 and .32, and [Schubert et al. \(2020\)](#), who find an elasticity between .24 and .37. My estimates split the difference between these ranges, which could be explained by my using market level variation in outside options, but estimating the effect

on individual wage growth through a wage renegotiation channel, rather than the change in local average wages. It could be the case that there is strategic complementarity in wage bargaining that causes wages to be more responsive to aggregate shocks rather than worker level shocks: firms are more likely to offer higher wages if they see a greater share of their workers credibly threatening to take another job.

Heterogeneous Effects

Top Destinations As in the mobility regressions, I estimate the effect of labor demand shocks for subsets of destinations: those within/across commuting zones, and Top 5/Non-Top 5 destinations. The specification that I estimate is similar: I instrument for subset-specific growth in nonlocal hires with a subset specific Network IV, and control for the complementary Network IV. Note however that the dependent variable is not destination specific.

Table 1.11 reports results. For destination with commuting zone overlap, I find that wage growth is somewhat less sensitive to shocks from MSAs within the same commuting zone (.166 versus .144 for MSAs that do have commuting zone overlap), though the difference is small and not statistically significant. Similarly, I find that wage growth is somewhat more responsive to nonlocal demand from non-top destinations, with an elasticity of .164 compared to .123 for shocks from top-5 destinations. The difference is slightly larger here, though again not statistically significant.

Effect by Education Level and Industry I also estimate heterogeneous effects on workers with different education levels and by their 2-digit NAICS sector of employment. Table 1.12 reports results.

For wage growth estimates, I find that the effect of nonlocal demand varies quite substantially by workers' education level. The strongest effect of nonlocal demand on wages is for workers with some college or an associate's degree with an elasticity of .227, substantially higher than the benchmark elasticity. For workers with a high school degree or less, the estimated elasticity is similar to the benchmark estimate, at .138 and .159. The effect of nonlocal demand for workers with a college education (Bachelor's degree or higher) is only

Table 1.11: Nonlocal Demand and Wage Growth: IV Estimates, Closely Connected Destinations

Dependent Variable: YoY Δ Log Wage		
Destinations within or Across Commuting Zones:		
	MSAs w/o CZ Overlap	MSAs w/ CZ Overlap
Δ Nonlocal Hires, Within CZ		0.144 (0.169)
Δ Nonlocal Hires, Across CZ	0.142 (0.0889)	0.166 (0.103)
N	178943	178943
Top-5 and Non-Top-5 Destinations		
Δ Nonlocal Hires, Top 5	0.123 (0.0868)	
Δ Nonlocal Hires, non-Top 5	0.164** (0.0668)	
N	179034	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is 2005-2019. Standard errors clustered at the MSA-by-education level. All specifications restrict to same-MSA and same-employer observations, and observations with absolute changes in hours less than or equal to 10. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects. Each row is a separate model that excludes the network IV for the denoted set of destinations, and controls for the complementary IV. J2J outflows are counted as within CZ if the destination MSA has any overlapping commuting zones with the origin. Top-5 destinations are the MSA-specific top 5 in the base year (2003).

.0625 and is not statistically significant.

Lastly, I estimate the effect of nonlocal demand on wages for workers in different 2-digit NAICS industries. I find that the strongest effect of nonlocal demand on wages is for workers in local service industries (retail trade; arts, entertainment, and recreation; and accommodation and food services) and manual industries (construction, manufacturing, wholesale trade, and transportation and warehousing) with elasticities of .139 and .135. For education and healthcare, I find an elasticity of similar magnitude that is not statistically significant. For other 2-digit NAICS industries, I find a null effect. One particular industry grouping to note here is the skilled scalable services (information; finance and insurance; professional, scientific, and technical; and management), which are characterized by urban agglomer-

ation and IT intensity, as well as workers with high education levels on average (Eckert et al., 2020).

Table 1.12: Nonlocal Demand and Earnings Growth: Heterogeneous Effects by Industry and Education Level

		Dependent Variable: YoY Δ Log Wage	
		IV	IV
		(1)	(2)
Δ Nonlocal Hires by Educ \times			
Education Level:		Industry:	
HS	0.159 (0.107)	Manual	0.135* (0.0708)
HS	0.138* (0.0829)	Skilled Scalable Services	0.0235 (0.0886)
Assoc & Some Coll	0.227*** (0.0860)	Healthcare & Education	0.131 (0.0847)
\geq College	0.0625 (0.110)	Local Services	0.139* (.0765)
N	179034		179034
First-stage F-stat	1069.2		1994.1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is 2005-2019. Standard errors clustered at the MSA-by-education level. All specifications restrict to same-MSA and same-employer observations, and observations with absolute changes in hours less than or equal to 10. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects. Industry groups shown are Manual (NAICS 23, 31-33, 42, 48-49), Skilled Scalable Services (NAICS 51, 52, 54, 55), Skilled Local Services (NAICS 61, 62), and Other Local Services (NAICS 44-45, 71, 72, 81). Omitted industries are NAICS 11, 21, 22, 53, 56, and 92.

1.6 Conclusion

There has been an increased focus on monopsony power in recent years, much of it studying on the conditions in local labor markets. While this research shows that local employment concentration has declined in the US in recent decades, a troubling trend is the decrease in migration and labor mobility. To the extent that job opportunities are nonlocal, this could offset the fall in local labor market concentration and perhaps even reverse the trend of declining monopsony power. My research adds to this by evaluating the relationship between geographic labor mobility and the nonlocal dimension of workers' job opportunities, which is crucial for evaluating the extent of monopsony power in the US.

Using data on job switches across metropolitan areas I show that nonlocal job opportunities are an important feature of US labor markets: about 40% of job switches are across, rather than within MSAs. Intuitively, job-to-job flows are a natural indicator of workers' outside options: when a worker switches jobs from one firm to another, this indicates a better job opportunity (or in the case of involuntary separations, a next-best opportunity). When aggregated, job-to-job outflows are a strong indicator of the opportunities available to workers outside of their local labor market. Longer distance job switches are relatively common: about half of those flows are across states, and more across commuting zones, suggesting that the non-local dimension of job flows is not simply a product of insufficient or incorrect geographic aggregation.

To evaluate the impact of nonlocal labor market opportunities on workers, I create measures of exposure to nonlocal labor market conditions based on historic rates of J2J flows across MSAs. I then estimate the effect of nonlocal labor demand shocks on labor mobility and on wage growth. When nonlocal labor demand is higher, I expect this to induce greater J2J outflows, as workers are more likely to encounter job opportunities outside of their local market. Furthermore, some workers may leverage outside offers to renegotiate wages at their current job rather than switching jobs, so I expect greater nonlocal labor demand to lead to increased wage growth even for workers that remain in the same job.

To estimate the effect of nonlocal labor demand on mobility and wage growth, I use an instrumental variable (IV) specification based on the Network Bartik IV approach introduced

by Schubert (2021). The instrument is a shift-share measure that uses historic J2J outflows to aggregate the effects of national industry shocks on labor demand in other MSAs. This specification addresses two major threats to causal inference by isolating variation due to changes in labor demand, rather than supply, and variation that is orthogonal to local labor demand. The key identifying assumption is the exogeneity of the initial J2J-outflow-weighted industry shares of nonlocal labor markets with respect to J2J outflows and wage growth.

I find that a typical shock to nonlocal labor demand increases J2J outflows by about 1.72 percent, and wage growth for job stayers by about 1.24 percent. Intuitively, job-to-job flows are more responsive to labor demand shocks from MSAs that have overlapping commuting zones, but labor demand shocks that are across commuting zones account for the majority of the overall response of labor mobility to nonlocal demand. However, labor mobility is much more responsive to demand from a small number of top destinations for each MSA, which account for about 80 percent of the total mobility effect. The estimates for wage growth show a somewhat different pattern: the effect of nonlocal demand is similar for MSAs regardless of overlap in commuting zones and for top and other destinations. I find limited evidence of heterogeneity in the mobility responses across education levels, but the effect of nonlocal outside options on wages is concentrated on workers without a 4-year college degree and in industries with lower average education levels.

These findings have several key implications. First, expanding workers' labor market opportunities through for example, remote work, can lead to greater productivity and wage growth by facilitating the reallocation of workers from low to high productivity areas or jobs. Importantly, these effects are mediated by the overall level of geographic labor mobility, which is influenced by several potentially offsetting secular trends: the overall decline in migration rates and labor market fluidity, the increase in longer-distance commuting, and the pandemic induced surge in remote work. There is potential for future research in this area to disentangle the contribution of these trends and cyclical factors in determining the extent that effects of local labor demand shocks are dispersed across regions. Second, expanding workers' labor market opportunities can lead to greater wages even conditional on productivity by improving workers' outside options and bargaining positions, which

speaks to a growing body of research on the extent of employer market power. The substantial heterogeneity in MSA-level J2J outflows suggests that geographic centrality is a potential confounder when estimating the degree of monopsony in local labor markets. While existing research has accounted for similar forms of cross-industry mobility and cross-occupation mobility, future research should consider geographic mobility as a source of workers' outside options.

1.7 Appendix to Chapter 1

Table A1.1: MSA Sample Summary Statistics

	N	Mean	Std. Dev.	P25	P75
J2J Outflow Rate	18415	0.0141	0.0071	0.0100	0.0165
Local J2J Rate	18415	0.0125	0.0068	0.0082	0.0151
Log YoY Employment	18415	0.0060	0.0284	-0.0050	0.0221
Log YoY Hires	18415	-0.0044	0.1288	-0.0467	0.0685
Log YoY Payroll	18415	0.0268	0.0412	0.0090	0.0451
Local Bartik	18415	-0.0012	0.1036	-0.0176	0.0526
Nonlocal Hiring Growth	18415	-0.0010	0.1172	-0.0181	0.0639
Network IV	18415	-0.0016	0.0992	-0.0142	0.0465
Network IV: Top-5 Destinations	18415	0.0000	0.0840	-0.0127	0.0443
Network IV: Non-Top Destinations	18415	-0.0010	0.0374	-0.0060	0.0193
Network IV: Within Commuting Zone	10284	-0.0005	0.0467	-0.0060	0.0162
Network IV: Across Commuting Zone	18355	-0.0016	0.0946	-0.0134	0.0492

Source: Job-to-Job statistics and Quarterly Workforce Indicators, U.S. Census Bureau. Sample period is 2005-2019. Observations are at the MSA level. Local Bartik is the local shift-share measure using 2003 industry composition and leave-1-out national industry growth rates. Nonlocal hiring growth is the sum of all nonlocal MSA YoY hires, weighted by 1-year lagged J2J outflow rates. Network IV is the sum of all nonlocal Bartik measures using leave-2-out national industry growth rates.

Table A1.2: CPS Sample Summary Statistics

	N	Mean	Std. Dev.	P25	P75
< High School	208806	0.04			
High School	208806	0.25			
Some College	208806	0.16			
Assoc	208806	0.11			
College	208806	0.28			
Grad	208806	0.15			
Male	208806	0.43			
White	208806	0.82			
Black	208806	0.10			
Asian	208806	0.06			
Hispanic	208806	0.13			
Age	208806	41.67	8.3	35	49
Hourly Worker	208806	0.44			
Weekly Earnings	208806	1109.13	656.39	620	1403.84
Hourly Wage (Reported)	92449	18.92	9.69	12.50	23.00
Hourly Wage (All)	208806	25.64	14.35	15	32.50
Usual Hours Worked	201864	42.78	6.46	40	43
Actual Hours Worked	208806	43.44	7.30	40	45
Log Hourly Wage Growth (Reported)	71057	0.043	0.19	0	0.091
Log Hourly Wage Growth (All)	208806	0.045	0.44	-0.036	0.140

Source: CPS-IPUMS, University of Minnesota. Sample period is 2005-2019. All statistics weighted by CPS earner study weights.

Table A1.3: MSA-Level Observations in CPS Sample

	MSA	MSA-by-Education
N	241	964
Mean	3232.37	1086.91
Min	42	2
P25	192	33
P50	370	79
P75	886	188.5
Max	14356	7589

Source: CPS-IPUMS, University of Minnesota. Sample period is 2005-2019. Number of MSAs and distribution of respondent observation counts in CPS extract used in wage regressions.

Table A1.4: First Stage Network IV Results

	Dependent Variable: Δ Log J2J Outflow Rate, Annual	
	IV: Hires (1)	IV: Hires & Employment (2)
Network IV: Hires	0.967*** (0.0223)	0.882*** (0.0226)
Network IV: Employment		0.836*** (0.0456)
N	73696	73696
R ²	0.244	0.260

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is quarterly data from 2005-2019. All models control for local changes in employment, hires, payroll and date and MSA-by-education-level fixed effects. Standard errors clustered at the MSA-by-education level.

Table A1.5: Robustness Check: Nonlocal Demand and Local J2J Flows, Top Destinations

	Dependent Variable: Δ Log Local J2J Rate	
	2SLS	IV
Δ Nonlocal Hires, Top 5	0.0482 (0.0350)	-0.0171 (0.0454)
Δ Nonlocal Hires, non-Top 5	0.0178 (0.0707)	-0.0348 (0.101)
N	19847	19847
Adj. R ²	0.0893	0.0891
First-stage F-stat	7064.2	1414.8

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is quarterly data from 2005-2019. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand. OLS* estimates use base-year (2003) exposure to nonlocal hires to instrument for 1-year lagged exposure to nonlocal demand. All models include MSA and date fixed effects and four quarterly lags of all controls and excluded instruments.

Table A1.6: Nonlocal Demand and Earnings Growth: Excluding Low-Observation MSAs

Observation Threshold:	Dependent Variable: YoY Δ Log Wage			
	MSA: P25 (1)	MSA: P50 (2)	MSA-Educ: P25 (3)	MSA-Educ: P50 (4)
Δ Nonlocal Hires by Educ	0.121* (0.0677)	0.126* (0.0663)	0.123 (0.0762)	0.121* (0.0724)
Local Bartik	0.333*** (0.123)	0.331*** (0.120)	0.353** (0.141)	0.229* (0.130)
N	178689	183050	164515	170999
First-stage F-stat	12431.9	13225.6	10130.5	11211.3

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is 2005-2019. Standard errors clustered at the MSA-by-education level. All specifications restrict to same-MSA and same-employer observations. Wage growth sample restricts to observations with absolute changes in hours less than or equal to 10. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects. Columns drop MSA or MSA-by-Education level bins with total observations less than the indicated percentile (Column 1: 192, Column 2: 370, Column 3: 33, Column 4: 79)

Table A1.7: Nonlocal Demand and Wage Growth: Nominal vs. Real Wage Growth Effects

	Nominal Wage Growth (1)	Real Wage Growth (2)
	Δ Nonlocal Hires by Educ	0.231*** (0.0772)
Local Bartik	0.379** (0.155)	0.361** (0.155)
N	121499	121499
First-stage F-stat	8486.7	8486.7

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Regional Price Parities (RPP), Bureau of Economic Analysis. MSA-level real wage growth adjustment is computed using the Bureau of Economic Analysis' Regional Price Parity (RPP) data. RPP sample period is 2008-2017. Standard errors clustered at the MSA-by-education level. All specifications restrict to same-MSA and same-employer observations. Wage growth sample restricts to observations with absolute changes in hours less than or equal to 10. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects.

Table A1.8: Nonlocal Demand and Earnings Growth: Job Stayers + Local Job Switchers

IV: Nonlocal YoY	Dependent Variable: YoY Δ Log Wage		
	Hires (1)	Cumulative Hires (2)	Cumulative Hires & Emp (3)
Δ Nonlocal Hires By Educ	0.0716 (0.0666)	0.0837 (0.0793)	0.0812 (0.0803)
Local Bartik: Hires	0.432*** (0.133)	0.432*** (0.133)	0.426*** (0.133)
Local Bartik: Employment			0.125 (0.141)
N	205105	205105	205105
First-stage F-stat	13544.0	5834.8	3399.1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is 2005-2019. Standard errors clustered at the MSA-by-education level. All specifications restrict to same-MSA and same-employer observations. Wage growth sample restricts to observations with absolute changes in hours less than or equal to 10. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects.

Table A1.9: Nonlocal Demand and J2J Outflows: No Local Controls

	Dependent Variable: Δ YoY Log J2J Outflow Rate	
	IV: Hires (1)	IV: Hires (2)
Δ Nonlocal Hires	0.306*** (0.0956)	0.339*** (0.0924)
Local Bartik: Hires	-0.0682 (0.0600)	-0.0364 (0.0669)
Δ Log Employment	-1.292*** (0.105)	
Δ Log Hires	0.0740*** (0.0234)	
Δ Log Payroll	0.0174 (0.0263)	
N	18415	18415
First-stage F-stat	641.8	580.1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is quarterly data from 2005-2019. All models control for local changes in employment, hires, payroll and date and MSA-by-education-level fixed effects. Standard errors clustered at the MSA-by-education level. IV estimates use a Network Bartik IV to instrument for 1-year lagged exposure to nonlocal demand.

Table A1.10: Nonlocal Demand and Wage Growth: No Local Controls

	IV: Nonlocal YoY Hires	
	(1)	(2)
Δ Nonlocal Hires by Educ	0.138** (0.0658)	0.139** (0.0613)
Local Bartik	0.435*** (0.139)	0.278** (0.132)
Lagged Network IV: Hires	-0.0469 (0.0678)	-0.0320 (0.0675)
Δ Local Employment, YoY	-0.289** (0.131)	
Δ Local Hires, YoY	0.255*** (0.0456)	
N	186075	186075
First-stage F-stat	14247.0	13591.7

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is 2008-2017. Standard errors clustered at the MSA-by-education level. All specifications restrict to same-MSA and same-employer observations. Wage growth sample restricts to observations with absolute changes in hours less than or equal to 10. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects.

Table A1.11: Nonlocal Demand and Wage Growth: No Recent Job Switches and Same Industry/Occupation Observations

	IV: Nonlocal YoY Hires		
	Benchmark (1)	Same Industry (2)	(3)
Δ Nonlocal Hires by Educ	0.138** (0.0658)	0.106 (0.0740)	0.118* (0.0689)
Local Bartik	0.435*** (0.139)	0.333** (0.157)	0.225 (0.149)
Lagged Network IV: Hires	-0.0469 (0.0678)	-0.0814 (0.0784)	-0.0558 (0.0783)
Δ Local Employment, YoY	-0.289** (0.131)	-0.172 (0.132)	
Δ Local Hires, YoY	0.255*** (0.0456)	0.232*** (0.0496)	
N	186075	147184	147184
First-stage F-stat	14247.0	13442.7	12435.1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Source: IPUMS-CPS, University of Minnesota. Quarterly Workforce Indicators and Job-to-Job statistics, U.S. Census Bureau. Sample period is 2008-2017. Standard errors clustered at the MSA-by-education level. Columns 2 and 3 restrict observations to respondents that report no job switches in the second CPS survey wave and work in the same industry and occupation in both wage observations. All models include controls for worker age, race, and education; 2-digit occupation and industry; and MSA and date fixed effects.

Chapter 2: Geographic Labor Mobility and On-the-Job Search

2.1 Introduction

Migration expands the set of job opportunities available to workers and the labor pool that firms can recruit from by integrating geographically distinct labor markets. Migration is also an important margin of adjustment in response to local economic shocks that can reduce workers' time to re-employment (Goetz, 2015), reduce persistence in local unemployment rates (Blanchard and Katz, 1992), or increase the local labor supply elasticity (Monte et al., 2018). However, migration has markedly declined since the 1980s (Molloy et al., 2016), and there is evidence that the migration response to local shocks has weakened (Dao et al., 2017). Moreover, recent work has studied the effects of variation in local economic conditions such as employment concentration (Azar et al., 2017; Rinz, 2018; Benmelech et al., 2018; Berger et al., 2022) or exposure to trade or other negative shocks (Autor et al., 2013; Mian et al., 2013), and there is evidence that the effects of such shocks are mediated by frictional connections to other markets (Adão et al., 2018). Understanding the role of local labor markets' geographic position is thus important in evaluating the opportunities that workers face and how they evolve in response to changing economic conditions locally or elsewhere.

In this paper, I document new facts on the prevalence of job-to-job flows across U.S. metro areas and introduce a general equilibrium model of spatial on-the-job search that provides a framework for studying geographically distinct labor markets with heterogeneous connections. Job-to-job flows are closely linked to workers' outside options, and their spatial dimension reflects the relative importance of the local labor market and other markets. I show that job-to-job flows across metro areas are a large component of all job-to-job flows,

which suggests that the spatial dimension of outside options is substantial. Furthermore, there is variation in metro areas' rates of job-to-job flows to and from other markets, and the direction of job-to-job flows is in line with the predictions of a gravity model, which points to heterogeneity in markets' geographic position in determining flows between markets.

The model that I develop replicates these features of the data and allows me to answer the following question: what is the effect of non-local job opportunities on labor market outcomes such as employment and wages? In answering this question, I focus two channels. First, by expanding workers' opportunities, access to non-local jobs improves workers' job ladders, increasing the pace of productivity enhancing reallocation and resulting in workers having better jobs in the aggregate. Second, when workers' outside options improve, even those that do not switch jobs can see increased earnings due to the improvement in their bargaining position (Caldwell and Harmon, 2019). In the model, wages are set through bargaining, so relaxing the frictions that limit mobility allows non-local job opportunities to affect firm rents.

Empirically, I find that spatial churn of workers is an important feature of the labor market and of job switching in particular. I use data on job-to-job flows within and across U.S. metro areas to show that job-to-job flows between metro areas are about 40% of all metro area job-to-job flows, and that gross cross-metro job-to-job flows are much larger than net flows. I also show that job switches across markets are important regardless level of geographic aggregation. While the majority of cross-metro job switches are within the same state, fewer than 12% are within the same commuting zone. Furthermore, long distance job-to-job flows are common: about 70% of cross-metro job switches are between cities that are more than 100 miles apart.

The aggregate statistics on the prevalence of cross-metro job-to-job flows mask heterogeneity in the rate that workers move to and from different metro areas, which points to the importance of heterogeneity in metro areas' geographic position. To study this, I estimate a classical gravity equation with market size and distance as explanatory variables. The gravity model explains a large share of the variation in bilateral flows, and the impact of distance on job-to-job flows is economically significant.

In order to understand the implications of heterogeneous connections between metro areas for labor market outcomes, I develop a general equilibrium model of spatial on-the-job search. The model combines features from [Schmutz and Sidibé \(2018\)](#) and [Shimer \(2006\)](#). [Schmutz and Sidibé \(2018\)](#) develop a partial equilibrium model of spatial on-the-job search which fixes the rate at which workers receive offers and the distribution of offered wages exogenously. The spatial environment of their model is defined by two bilateral frictions: moving costs and search frictions. These frictions reduce the effective size of the labor market that workers face by limiting the workers' ability to receive and accept job offers from other regions. I extend this model by introducing general equilibrium features. First, I introduce a spatial formulation of the matching function, which allows me to endogenize the rate at which workers receive job offers from firms in different markets. Second, I endogenize firms' vacancy posting decisions. Third, I endogenize the wage distribution by introducing strategic bargaining over wages. Because standard bargaining solutions such as surplus-sharing or Nash bargaining are not generally applicable to models with on-the-job search, I implement the bargaining solution from [Shimer \(2006\)](#). These features allow me to study the evolution of job-to-job flows and the geographic distribution of market tightness in general equilibrium. Because workers relocate and firms change their vacancy posting decisions in response to local shocks, accounting for endogenous market tightness and equilibrium offer rates across different markets will be critical for determining the long run responses of outcomes such as wages, rents, and mobility to local shocks.

I calibrate and solve a simple version of the model with two symmetric cities. The calibrated moving costs are much lower than standard estimates in the literature, consistent with the finding in [Schmutz and Sidibé \(2018\)](#) that including search frictions reduces estimated moving costs. I provide a novel explanation for this result: in a search framework, the premium that workers require to move is decreasing with respect to workers' initial wages, since workers lower on the local job ladder expect to encounter better jobs locally. Estimates of moving costs based on present value calculations that do not take this channel into account will tend to overstate moving costs.

Lastly, I use the calibrated frictions to study the contribution of frictional connections between markets to labor market outcomes. Despite there being no potential for spatial mis-

allocation in the symmetric city model, I find that relative to an economy with no labor mobility across cities, average wages are about 6% higher and the unemployment rate is 18% lower. These gains arise from the interaction of geographic mobility and search frictions: improving matching across geographically distinct markets reduces the effective degree of search frictions in the economy, leading to lower frictional unemployment and a higher rate of switching to more productive jobs. Because markdowns of wages relative to workers' marginal product are an endogenous outcome of the model, I can decompose the change in average wages into two channels. First, by increasing offer rates, connections to other markets strengthen job ladders and improve the productivity distribution of jobs. Second, connections to other markets improve workers' bargaining position and result in higher wages conditional on the productivity of the job. I find that this second channel is quantitatively significant: lower markdowns account for about half of the difference in average wages relative to an economy with no labor mobility.

2.2 Related Literature

In this paper I present new facts on the rate of job-to-job flows across U.S. metro areas, which contributes to a recent literature that has documented trends in internal migration and labor mobility. [Kaplan and Schulhofer-Wohl \(2017\)](#) shows that the annual gross rate of interstate migration has declined from 3% in 1991 to about 1.5% in 2011. Migration rates are declining across different demographic groups, ruling out an aging population or other demographic changes as an explanation, and suggesting that the decline is consistent with declining incentives to migrate. Incentives to migrate have declined because job opportunities in different areas becoming more similar over time, and because improved information has reduced the need to move to learn about other locations. They find that these mechanisms can explain roughly half of the decline in annual migration rates. [Molloy et al. \(2014\)](#) document similar declines in within-county and across-county migration. They use survey responses from the Current Population Survey (CPS) to show that job-related moves are a major component of long-distance migration and that geographic mobility and labor market transitions are strongly correlated. They suggest that declining migration rates are linked to the concurrent decline in labor market fluidity ([Davis and Haltiwanger, 2011](#); [Hyatt and Spletzer, 2013](#)). [Molloy et al. \(2016\)](#) investigates potential causes of the decline in

migration and labor market fluidity, finding some role for demographic change and potentially for changes in the firm size and age distribution and declining social trust. [Molloy and Smith \(2019\)](#) present updated trends, showing that migration rates have stabilized since the end of the Great Recession, which suggests that migration may be pro-cyclical. They also document patterns in the direction of migration. They find that metro areas with stronger labor demand have higher outflow rates as well as inflow rates, which suggests that churning of workers through different labor markets is important. They also find that out-migration from metro areas with high labor demand is directed towards other high-demand areas, and that out-migration from weak areas is more likely to be to another weak area. I hope to document similar facts using the job-to-job flows data and provide theoretical justification for such a core-periphery structure.

This paper is closely related to the literature studying the geographic frictions that limit worker mobility. A common finding in this literature is that high moving costs are needed to justify observed labor mobility. For example, [Kennan and Walker \(2011\)](#) estimate a dynamic model of migration and find that moving costs in the range of \$300,000 are needed to justify observed mobility given the present value of expected earnings differences across locations. [Bartik \(2017\)](#) generalizes mobility costs to include sector and occupation switching as well as geographic relocation and estimates these costs in a model of worker job choice, finding moving costs in the hundreds of thousands of dollars. From the trade literature, [Artuc et al. \(2007\)](#) and [Artuç and McLaren \(2015\)](#) find that similarly high mobility costs are needed to justify observed changes in geographic and occupation shares of employment in response to trade shocks. Taking a different approach, [Kosar et al. \(2019\)](#) use hypothetical choice questions in the New York Fed's Survey of Consumer Expectations to estimate willingness-to-pay for location and mobility options. They find that on average, there are high non-pecuniary costs to moving - over 100% of annual income - in addition to strong preferences for locations near family members and local amenities such as social norms. However, they do find substantial heterogeneity in how people classify themselves in terms of mobility and in the non-pecuniary moving costs that different people face. Some face very low moving costs while others are "rooted" and face very high moving costs.

The paper that is most closely related to mine is [Schmutz and Sidibé \(2018\)](#), who develop

a spatial on-the-job search model and jointly estimate bilateral moving costs and search frictions between French cities. Their model is a generalization of the canonical wage-posting model by [Burdett and Mortensen \(1998\)](#) with multiple markets. Employed workers continually receive job offers from other firms in different locations at a fixed rate. While employed and unemployed workers are free to move to a new location without a job offer, due to congestion externalities via endogenous location-specific amenities ensure that labor mobility of the employed only takes place when workers receive a job offer that induces a move. Accounting for the stochastic nature of migration, this reduces the implied moving costs necessary to explain observed mobility. They estimate their model using data on job transitions from a French matched employee-employer dataset, using the rate of job-to-job flows across markets to identify the bilateral offer rate, while using the average difference in wages for movers to identify moving costs. They find average moving costs that are an order of magnitude lower than other estimates in the literature, roughly 15,000 Euros.

I build on this work in two ways. First, I generalize the random, spatial search framework of [Schmutz and Sidibé \(2018\)](#) to include endogenous vacancy posting and wage bargaining. In my model, market tightness and job finding rates in each city are endogenous, but linked to underlying search frictions. Second, I intend to estimate the model using U.S. data on job-to-job flows across metro areas, which will allow me to compare moving costs estimates to those in the literature studying the U.S. that do not take into account the contribution of search frictions in limiting geographic mobility.

A number of other papers study the interaction of search frictions and migration in general equilibrium. [Karahan and Rhee \(2014\)](#) study the indirect effects of population aging on migration in a search context with bargaining. In their model, firms can direct their recruitment efforts locally or towards other markets. Older workers face higher moving costs and have weaker outside options, so firms can extract higher rents from them relative to younger workers. When older workers make up a larger share of the labor force, firms are more incentivized to recruit locally, which reduces migration. They estimate that the indirect effect of aging accounts for about half of the decline in cross-state migration since the 1980s. [Zhang \(2017\)](#) uses a two-region wage posting to study the spillover effects of minimum wage increases across nearby counties, taking into account labor mobility and firms'

vacancy posting decisions. Unlike [Schmutz and Sidibé \(2018\)](#) or the model I develop, workers and firms in either city are equally likely to encounter each other. Lastly, [Schluter and Willems \(2018\)](#) integrates directed search and the random-utility spatial equilibrium model in a many-city context with worker heterogeneity. Random utility over locations helps avoid unidirectional migration flows since workers will allocate their search efforts to different locations based on their preference shocks. Their model also accommodates bilateral search frictions (in the form of search costs) and moving costs, but the directed search environment does not have an associated matching function.

The model I introduce also contributes to the literature on local labor markets and spatial equilibrium by introducing endogenous market tightness as a force for dispersion of economic activity in spatial equilibrium. In the canonical spatial equilibrium model of [Roback \(1982\)](#), labor is perfectly mobile and the housing supply is inelastic. In this case, workers are unaffected by shocks to local labor demand or amenities, because such shocks are fully capitalized into housing prices. When workers have idiosyncratic preferences for locations as in [Moretti \(2010\)](#), preferences act as a force for dispersion that limits the mobility response to local shocks. While marginal workers are indifferent across cities, in this setting, inframarginal workers can be affected by local shocks even with a fully inelastic housing supply. More generally, the welfare implications of local productivity or other labor demand shocks in this environment depend on whether or not the local labor supply or housing supply is more elastic. [Hsieh and Moretti \(2015\)](#) applies this framework to data on local wages and employment, finding that growing dispersion of wages across cities is closely related to growing housing cost dispersion, and that the spatial misallocation induced by housing supply restrictions in productive cities is a substantial drag on U.S. output.

Lastly, I contribute to the literature that studies the response of labor markets to local shocks. The seminal paper in this literature is [Blanchard and Katz \(1992\)](#), which shows that shocks to the local unemployment rate are transitory and disappear after about a decade, and that labor mobility is an important margin of adjustment. More recently, [Dao et al. \(2017\)](#) shows that net migration is strongly counter-cyclical, and responds more strongly to unemployment disparities across regions during recessions. However, they find that the

out-migration response from areas that receive negative shocks has weakened since the 1990s. Another series of papers studies the spatial dynamics of unemployment. [Goetz \(2015\)](#) uses worker-level data from LEHD to show that moving is associated with longer unemployment spells in reduced form regressions, but that instrumenting for migration using local house prices implies that mobility actually reduces the time to re-employment. This suggests that migration is a response of those who face particularly poor labor market conditions and have strong incentives to relocate. [Şahin et al. \(2014\)](#) study the spatial mismatch of unemployment and vacancies during the Great Recession. In a strictly local model of the labor market, differing locations of vacancies and the unemployed creates mismatch unemployment, since the job finding rate would be higher if the unemployed were in the same place as vacancies. However, they find that the contribution of mismatch to unemployment in the Great Recession is small. Lastly, [Bilal \(2019\)](#) studies the determinants of local unemployment rates in spatial equilibrium. As in the model I present, [Bilal \(2019\)](#) studies firms' decisions, and shows that firms have more incentive to locate near other productive firms, which leads to high unemployment in more remote areas. I contribute to this literature by deriving a spatial matching function that allows workers and vacancies in any market to be matched with each other.

2.3 Motivating Facts

In this section, I use data on quarterly cross-metro job-to-job flows in the US to show that geographic mobility is an important feature of the US labor market ¹.

2.3.1 Data: Job-to-Job Origin-Destination Statistics

This section makes use of the Job-to-Job Origin-Destination statistics (J2JOD) from the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) database. The LEHD is a comprehensive matched employee-employer database with detailed employment histories of workers based on state-level unemployment insurance records. The J2J statistics are public-use tabulations of job switches in the LEHD, including the J2JOD

¹Chapters 1 and 2 of this dissertation share a common motivation: understanding the role of geographic labor mobility in the US labor market. As such, material in this section is closely related to the descriptive statistics presented in Section 1.3, and in some cases overlaps.

statistics which track the number of job-to-job flows across industry and geography pairs. For further details on the LEHD and the J2JOD statistics, see section 1.3.1.

For my analysis, I use metro level data on job-to-job transitions and wages. Using metro-level rather than state-level flows greatly increases the dimensionality of the data, but metro areas are a reasonable approximation of local labor markets which will allow me to study cross-market linkages. Aggregating up to the metro level obscures some important details regarding cross-industry flows and worker heterogeneity, but provides sufficient geographic detail to map to the model I propose. I combine both within- and adjacent-quarter J2J flows in the measures that I use. Data on average earnings for job switchers is based on full-quarter earnings and is restricted to flows between stable jobs. A stable job is defined as employment that lasts for at least one full quarter. For this to be the case, employment must be observed for at least three consecutive quarters. A stable job-to-job flow occurs when a worker transitions from one stable job to another. Average earnings for the previous and current job are reported in the J2JOD for stable job-to-job flows using the closest full quarter of earnings for each job. For within-quarter job-to-job flows that take place in period t , this means that average earnings are reported for workers in period $t - 1$ and period $t + 1$.

The other main source of data is the Quarterly Workforce Indicators (QWI), which are also constructed from the LEHD, and report gross hires and separations, average earnings, and other labor market statistics for locations and industries. I use metro-level average earnings, employment, and several measures of labor market fluidity and business dynamism constructed from hires and separations in the QWI.

There are several caveats about these data sources. The QWI and the J2JOD are both constructed from noise-infused microdata for disclosure protection purposes. For particularly small cells, the values for job-to-job flows are synthetic, though true zeros are preserved in both the noise-infusion and synthesizer. This introduces measurement error, but I do not expect aggregate statistics to be affected. In addition, while the J2JOD is well-suited for studying aggregate patterns in geographic job-to-job flows as well as the determinants of bilateral job-to-job flows, it does paint only a partial picture of labor mobility since moves that start or end in unemployment are not considered as job-to-job flows. Furthermore, while the J2JOD has average earnings of job switchers in the origin and destination loca-

tions, results later in this paper show that the average gain of movers can be misleading, since the premium over current wages that workers' require to move depends on their initial position in the local job ladder.

2.3.2 Sample

To ensure that statistics are consistent over time, I construct a balanced panel using metro areas that are available from 2003q1 to 2017q4. This preserves about 90% of the metro pair-quarter observations in the full sample, and the excluded metro areas are generally small. There are two notable exclusions: the Washington-Arlington-Alexandria metro area, which does not enter the LEHD until 2005, and the Boston-Cambridge-Newton metro area, which does not enter until 2010. Because job-to-job flows are highly seasonal, I annualize all series by taking the average of the quarterly values for each year.

2.3.3 Descriptive Analysis

The model of spatial on-the-job search that I introduce below generates specific predictions about job-to-job hires across different labor markets that I hope to validate with the J2JOD data. First, the model predicts that job-to-job flows across different markets are important. If job-to-job flows across metro-areas are instead negligible relative to within-metro flows, then a model of strictly local labor markets sufficiently characterizes the data. Second, the dynamic and stochastic nature of migration resulting from different opportunities that individuals encounter in the model means that the gross job-to-job flows should dominate the net job-to-job flows. If net migration rather than churn is instead the dominant feature of the data, then a model of search within cities and static migration across cities is sufficient to characterize the data. Lastly, heterogeneous spatial connections between cities in the model imply that there should be different rates of job-to-job inflows and outflows in different metro areas. Specifically, metro areas that are closer to other large metro areas should have a higher degree of churn. I investigate these predictions in turn.

2.3.4 How important are Cross-Metro Job-to-Job Flows?

The first question I address using this data is the importance of labor mobility across metro areas. Table 2.1 lists within-metro and across-metro job-to-job flow rates as a percentage of employment, hires, and all job-to-job flows. The average quarterly job-to-job flow rate is just over 4% of employment and job-to-job hires account for 26% of all hires. On average, cross-metro job-to-job hires make up about 39% percent of total job-to-job hires in metro areas, or 10% of total hires.² Overall, this suggests an important role that interactions between geographically distinct labor markets play in shaping workers’ job opportunities and firms’ recruitment pools.

Table 2.1: Within-Metro and Across-Metro Job-to-Job Flow Rates

% of	All J2J	Within-Metro	Across-Metro		
			Total	Net	Excess
Employment	4.20	2.57	1.63	0.09	1.54
Hires	26.84	16.44	10.40	0.55	9.85
All J2J	100	61.27	38.73	2.05	36.68

Average quarterly job-to-job flows over 2003-2017. Net flows is the sum of the absolute value of net cross-metro job-to-job hires in each city. Excess flows is the difference between gross cross-metro J2J and net cross-metro J2J. Source: J2JOD, U.S. Census Bureau

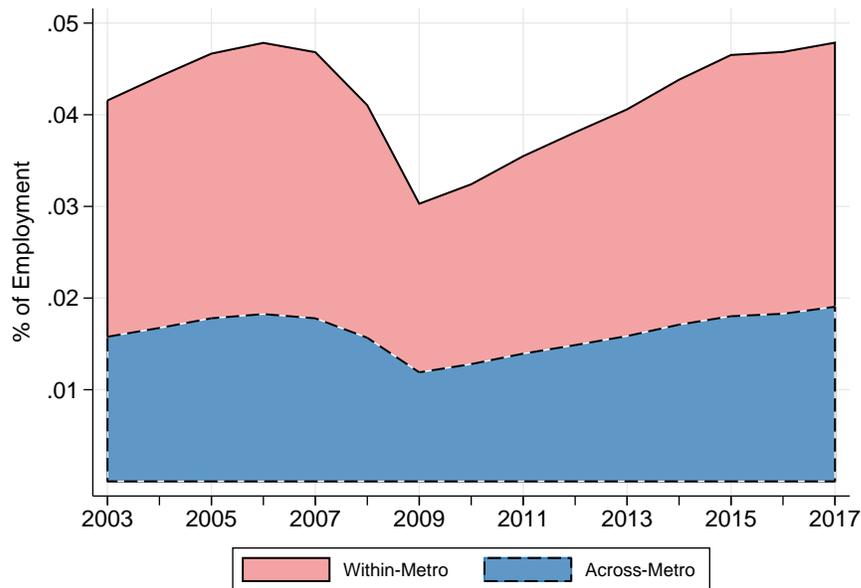
One of the main features of the theory I propose is that migration is a steady state phenomenon. If we see that migration is dominated by churn and not net flows, this suggests that mobility is mostly due to idiosyncratic opportunities that individuals face, rather than structural shifts in the size of local labor markets.

Referring again to Table 2.1, the last two columns decompose cross-metro job-to-job flows into net flows, defined as the sum of the absolute value of net cross-metro J2J hires for each metro, and excess flows, or the reallocation of workers across metro areas above what is needed to account for net reallocation. I find that this excess reallocation or churn accounts for most cross-metro job switches: excluding net flows reduces the cross-metro share of job-to-job hires only from 39% to 37%.

²Note that this is not the share of total job-to-job flows in the U.S. because the sample is restricted to job-to-job flows within or across metro areas. Job-to-job hires that originate from or end in a location outside of a metro area are not included in this count. Rather, this tells us the share of job-to-job flows starting and ending in a metro area that end in a metro different than the origin. Because this includes all within-metro flows but not all flows to or from each metro, this figure likely understates the extent of labor mobility that is relevant for metro areas

To see trends in cross-metro job-to-job flows, Figure 2.1 plots within-metro and across-metro J2J rates as a percentage of employment over time. The total J2J rate is highly cyclical, declining steeply during the Great Recession. The J2J rate recovered post-2010 and in recent years has slightly surpassed its pre-recession peak. The cross-metro J2J rate is strongly correlated with the aggregate rate, falling during recession and showing a surprisingly strong recovery. Cross-metro flows are somewhat less cyclically sensitive than within-metro flows, so their share of total J2J saw a slight increase during the Great Recession.

Figure 2.1: Within-Metro and Cross-Metro J2J Rate



The top line is total annual job-to-job flows as a percentage of employment. The upper shaded area is the contribution of within-metro flows, and the lower shaded area is the contribution of cross-metro job-to-job flows.

Geographic Dispersion

Table 2.2 shows the share of cross-metro job-to-job hires across different distances and geographic boundaries. While distance does play a role in influencing the number of J2J flows, longer-distance job switches are fairly common: almost 40% of switches take place between MSAs with a distance of 250 miles or greater. Most cross-metro J2J hires are within state, about 58%. However, this leaves a substantial share of flows that are across

Distance: Less Than		Within:	
25 miles	10.87%	Commuting Zone:	
50 miles	20.98%	Lower Bound	6.43%
100 miles	32.75%	Upper Bound	11.59%
250 miles	65.01%	State	58.12%
500 miles	90.00%	Division	72.40%
1000 miles	99.13%	Region	80.11%

Share of average quarterly job-to-job flows across metro areas over 2003-2017. Lower bound for CZ includes flows between single-CZ MSAs. Upper bound includes flows between MSAs with any overlap in CZ. Source: J2JOD, U.S. Census Bureau

states. To decompose these I also show the share of flows that are across the nine Census divisions and four Census regions.³ About 20% of cross-metro flows are across divisions or regions.

I also check the number of cross-metro job switches that are potentially within the same commuting zone. Commuting zones are a data-driven aggregation of counties based on commuting flows, which makes them a prime candidate for an alternative labor market definition. There is no strict hierarchy between CBSAs and commuting zones, so I cannot obtain an exact estimate of cross-metro, within CZ flows, but I can establish a range of potential figures. For a lower bound, I count a cross-metro flow as within-CZ if the two metro areas contain only one commuting zone that is common between them. For an upper bound, I count any flow between metro areas with any overlap in commuting zones as within-CZ. At most, 11.5% of cross-metro flows are within the same commuting zone. This is consistent with the findings that job switches across metro areas are geographically dispersed and that long distance moves are fairly common.

2.3.5 MSA-level Heterogeneity

Lastly, the framework I propose emphasizes heterogeneous spatial linkages between markets, which imply that different metro areas should have different rates of job-to-job churn. More connected places should have higher inflows and outflows, and remote places should have fewer workers moving through them. Other metro level characteristics may also play

³The four Census regions are Northeast, South, Midwest, and West. Each region is further divided into divisions.

a role in the share of job-to-job flows that are non-local. For example, in larger metro areas J2J outflows tend to comprise a smaller share of all job-to-job separations since the local labor market is relatively more important⁴. To see the role of geographic heterogeneity, Table 2.3 shows the metro areas that have the lowest and highest share of non-local J2J separations, conditional on metro area employment.

Table 2.3: J2J Outflows as a share of all J2J Separations

Bottom	MSA	Non-local J2J Share	Employment	Residual
1	Billings, MT	0.29	78,731	-0.319
2	Burlington, VT	0.27	111,560	-0.310
3	Bismarck, ND	0.36	62,521	-0.269
Top				
3	Riverside-San Bernardino, CA	0.61	1,198,300	0.24
2	Vallejo-Fairfield, CA	0.84	138,940	0.28
1	Trenton, NJ	0.82	227,276	0.30

Average share of cross-metro J2J separations in overall J2J separations. Top/Bottom based on residuals from regression of non-local J2J share on the log of employment

The bottom three metro areas by this measure have non-local job-to-job quit shares around 30%. They are small MSAs, and size alone would predict that their non-local quit share would be close to 60%. However, these places are notably distant from other population centers of the US, being in Montana, Vermont, and North Dakota. The metro areas near the top of this measure have non-local quit shares between 60% and 80%, meaning a majority of job-to-job quits are to another metro area. For these places, local employers comprise only a small share of the job opportunities that workers actually face. These metro areas close to other, larger MSAs. Trenton, New Jersey is between the Philadelphia-Camden and New York City metro areas, Vallejo-Fairfield is in the San Francisco Bay area, and Riverside is near Los Angeles. At least qualitatively, this evidence suggests that heterogeneity in metro areas' geographic positions affects the job opportunities that workers face.

To see if this pattern holds moer broadly, I estimate a classical gravity equation based on size and the distance between markets and show that it provides a reasonable characterization of job-to-job flows across metro areas.

⁴J2J outflows are share job-to-job separations originating in a given location that are associated with a hire in a different location.

In the classical gravity framework, bilateral trade flows are proportional to the product of the size of the two economies in question and the distance between them. While useful in a descriptive sense, the downfall of the classical gravity framework is that it is purely bilateral, meaning that it ignores multilateral resistance, or the diversion effects that other countries can have on bilateral trade (Head and Mayer, 2014). In the context of labor mobility, the gravity framework has also proved useful for characterizing flows of migrants across countries (Beine et al., 2016). The canonical micro-foundation for a gravity framework in migration is a random utility model, in which shocks to idiosyncratic preferences combined with bilateral costs or other determinants of migration generate a structural equation, in which origin and destination fixed effects control for multilateral resistance. To evaluate how well a classical gravity framework fits the J2J data, I use the balanced panel of U.S. metro areas described in section 3.2 to estimate the following equation:

$$\log(J2J_{ckt}) = \beta_0 + \beta_1 \log(EMP_{ct}) + \beta_2 \log(EMP_{kt}) + \beta_3 \log(Dist_{ck}) + \beta_4 Comp_{ck} + \delta_t + u_{ckt} \quad (2.1)$$

Where $J2J_{ckt}$ is the number of job-to-job hires in city c originating in k in period t , EMP_{kt} is total quarterly employment in the origin or destination, $Dist_{ck}$ is the distance between metro areas, and $Comp_{ck}$ is a measure of the dissimilarity of the employment composition between metro areas. I lag employment four quarters in order to avoid mechanical correlations between job-to-job separations and hires and total employment. To calculate distance I use the coordinates for each metro area that are given in the Census delineation files, and take the geodesic distance between each coordinate pair. Because metro areas are different sizes and the geographic center of each metro area may not correspond to the core of the population weighted center, this distance is likely measured with error. However I expect this measure to be a suitable proxy for distance-related frictions to mobility. Lastly, for employment dissimilarity I use the Euclidean distance between the vectors of employment shares of 3-digit NAICS industries.⁵

⁵To construct the measure of employment composition distance, I use 3-digit NAICS employment shares in each metro from the QWI. For each MSA pair, the employment composition distance is the sum of squared differences between employment shares in each industry.

Table 2.4: Gravity Regression, J2J Flows to Destination Metro

	(1)	(2)	(3)	(4)	(5)
Log Employment		0.418***	0.502***	0.497***	0.716***
Log Origin Employment		0.409***	0.494***	0.489***	0.845***
Log Distance			-0.823***	-0.819***	-1.034***
Employment Composition Distance				-0.577***	-5.599***
Date FE	yes	yes	yes	yes	yes
Origin & Dest. FE	no	no	no	no	yes
N	1433981	1194068	1194068	1194068	1194068
Adj. R ²	0.00153	0.260	0.544	0.545	0.670

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Balanced panel starting in 2003. Standard errors clustered within destination metro. Employment is lagged one year. Employment distance is the euclidean distance between the vector of 3 digit NAICS employment shares in the origin and destination.

Table 4 reports estimates of the classical gravity equation. Because I am interested in model fit, the improvement in R^2 from including size and distance is an important metric, so the first column regresses the log of job-to-job flows on the time fixed effects only. The R^2 in this specification is close to zero. Including the total employment of the origin and destination increases the R^2 to about .26. Including the log of distance in column (3) raises the R^2 to .544, which indicates that distance is an important determinant of the linkages between labor markets. The distance elasticity is roughly equal to -.8, so conditional on their size, metro areas that are 10% farther away from each other have roughly 8% fewer job-to-job flows between them. The employment elasticity, β_2 , is similar in magnitude for both the origin and destination, at about .5 for each. The coefficient for employment composition distance is negative, implying that places with less similar industrial composition have fewer job-to-job flows between them. However, the effect on model fit is small. Overall, the high R^2 and the strong distance elasticity suggests that a gravity model fits the data on cross-metro job-to-job flows fairly well.

Lastly, to explore other geographic frictions, I estimate the effect of being in different states or commuting zones. Results are reported in Table 5. I find that conditional on distance, being in a different state or commuting zone reduces the number of job-to-job flows quite substantially. This suggests that physical distance only partly determines the frictional connections between labor markets.

Table 2.5: Gravity Regression, J2J Flows to Destination Metro

	(1)	(2)	(3)
Log Employment	0.716***	0.698***	0.638***
Log Origin Employment	0.845***	0.827***	0.778***
Log Distance	-1.034***	-1.024***	-0.674***
Employment Composition Distance	-5.599***	-5.327***	-4.364***
Cross Commuting Zone		-0.537***	
Cross State			-2.254***
Date FE	yes	yes	yes
Origin & Dest. FE	yes	yes	yes
N	1194068	1194068	1194068
Adj. R2	0.67	0.674	0.772

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Balanced panel starting in 2003. Standard errors are clustered within destination metro. Employment is lagged one year. Employment distance is the euclidean distance between the vector of 3 digit NAICS employment shares in the origin and destination.

2.3.6 Recap

In this section, I have shown that job-to-job flows across metro areas are a substantial share of all job-to-job flows. To the extent that job switching reflects job opportunities that workers face, this suggests that non-local jobs are an important component of workers' options. Furthermore, I show that job-to-job flows across MSAs are dominated by churn rather than net migration, and that there is heterogeneity both in bilateral flows and in the overall level of churn in metro areas that is related to their geographic position.

The model of spatial on-the-job search that I present in the next section replicates these stylized facts. By expanding the standard job ladder framework to allow workers and firms in different locations to match with each other, I introduce job-to-job flows across metro areas. In steady state, all migration in the model is due to churn. Furthermore, frictional connections between markets are the source of geographic heterogeneity that we see in the gravity estimates. The model provides a framework for studying the implications of these geographic frictions on labor market outcomes, and particular will allow me to study the effects of non-local job opportunities on wages and employment.

2.4 Model: Spatial On-the-Job Search

In this section, I develop a general equilibrium model of spatial on-the-job search that provides a structural foundation for understanding job-to-job flows across geographic areas. The model accounts for workers' mobility decisions, firms' vacancy posting decisions, and their strategic bargaining over wages. This provides two advantages over existing models of search and migration. First, by introducing a role for vacancies, the model provides a microfoundation for reduced-form analysis of cross-metro job-to-job flows using the spatial matching function. Second, by explicitly modeling labor demand and the wage bargaining process, the model allows for analysis of the evolution of workers' outside options and labor mobility in response to regional shocks.

2.4.1 Environment

There is a set of N labor markets, which I refer to as cities but may be interpreted as cities or city-industry pairs in different contexts. Each market, c , is differentiated by its local distribution of productivity z , given by $H_c(z)$, and by their spatial connections to other labor markets. The spatial connections take the form of two frictions: bilateral search effectiveness a_{ck} , which affects the rate that workers in market k encounter firms in market c , and bilateral moving costs γ_{ck} . Time is continuous.

Workers

While worker heterogeneity, particularly by age and education, is surely important in determining search behavior, job ladders, and mobility, in the benchmark model, I assume that workers are homogeneous in all aspects except their current employment status and location. Workers continually search for better jobs, but are not constrained to search in their own market. Workers face moving costs to relocate to employment or unemployment in a new city. Because of congestion externalities in search, only currently unemployed workers will choose to move to unemployment in another city in equilibrium.

Firms

Single-worker firms produce a single good using CRS technology, taking the output price as given. The price of the good is fixed and constant across all markets. There is a continuum of ex-ante homogeneous firms that can choose to enter the labor market by posting a vacancy. Upon matching with a worker, both parties observe the productivity of the match, which is drawn from a fixed, exogenous distribution specific to each location, $H_c(z)$.

After matching and observing the productivity of the match, firms and workers bargain over wages.

Labor Market

The labor market is subject to search and matching frictions. Both unemployed and employed workers continually search for jobs. Search is random. In addition to job-to-job quits, there is an exogenous probability of job destruction that is constant across all labor markets.

Unemployed workers have a constant search intensity equal to one, so the mass of searchers from unemployment in city c is simply n_c^u . Let n_c^e be the number of employed workers in city c . The employed search with constant intensity $\sigma \in (0, 1]$, so the mass of searchers from employment is σn_c^e . The total number of searchers in city c is $s_c = n_c^u + \sigma n_c^e$.

2.4.2 Spatial Matching Function

To generalize the standard matching framework to accommodate cross-market matches, I introduce a spatial matching function. This is a novel feature of my model that will allow for endogenous cross-market offer rates. Vacancies posted by firms in city c can be matched with local searchers or those in other cities, and workers have different likelihoods of being matched with a job in city c based on their residence. The likelihood is determined by bilateral search effectiveness a_{ck} , which determines the effective number of searchers from city k that enter into the matching function for city c .⁶ The effective mass of searchers

⁶At this point I place no restrictions on within-city search effectiveness, a_{cc} , but in the benchmark calibration I normalize a_{cc} to one.

for city c is then simply the weighted sum of all searchers, which enters into the matching function to determine the total number of offers

$$s_c^* = \sum_{k=1}^N a_{ck} s_k \quad (2.2)$$

$$y_c = Am(v_c, s_c^*) \quad (2.3)$$

Equation (2.3) determines the total number of matches made in city c , y_c . This depends on aggregate matching efficiency A and a CRS matching function $m(v, s)$, where v_c is the total number of vacancies posted by firms in city c and s_c^* is the effective mass of searchers in city c . The generalization I make is straightforward: the effective mass of searchers is a weighted sum of job searchers in all markets, where the weights are given by bilateral search effectiveness a_{ck} .

Because matching is random, the number of bilateral offers from firms in city c to workers in k depends on the latter's effective share of s_c^* :

$$y_{ck} = \frac{a_{ck} s_k}{s_c^*} y_c \quad (2.4)$$

We can further define the matching probabilities of searchers and firms. Unemployed workers in k match with firms in c at the rate ϕ_{ck}^U , and employed workers match at rate ϕ_{ck}^E . Note that these two rates can differ because the search intensity of the employed can be less than the search intensity of the unemployed. Firms in c match with workers in k at rate ψ_{ck} .

Unemployed:	$\phi_{ck}^U = y_{ck} / s_k$
Employed:	$\phi_{ck}^E = \sigma y_{ck} / s_k$
Firms:	$\psi_{ck} = y_{ck} / v_c$

2.4.3 Workers

The equilibrium is characterized by firms' and workers' value functions over vacancies and employment states. In contrast to a standard search model, value functions are indexed by location.

Wages will ultimately be set by bargaining, but for the workers' problem, it is convenient to write value functions in terms of the local wage distribution, $F_k(w)$. In equilibrium, the wage distribution will be determined by the exogenous productivity distribution of vacancies and the outcome of the bargaining, but workers take the decisions of firms and the bargaining outcome as given.

Let U_k be the value of being unemployed in city k , and $E_k(w)$ be the value of being employed in market k in a job that pays wage w . As in, [Schmutz and Sidibé \(2018\)](#) I assume that workers in either state consume a size-dependent local amenity ξ_k , which can be seen as a proxy for an inelastic housing supply that absorbs variation in productivity and wages across locations. The value of unemployment is equal to the flow unemployment benefit b plus the expected value of the worker's employment state in the future. Unemployed workers in k receive offers from other markets at rate ϕ_{ck}^U and decide whether or not to accept these offers. Time is continuous, so I rule out workers receiving multiple job offers simultaneously. For employed workers, the flow value of employment is the wage w . At exogenous rate δ , the job is destroyed and the worker transitions to unemployment. Workers that remain employed receive offers from other markets at rate ϕ_{ck}^E and decide whether or not to accept them and pay the mobility cost γ_{ck} .

$$rU_k = b + \xi_k \sum_c \phi_{c,k}^U \int_{\underline{w}}^{\bar{w}} \max \{E_c(w) - \gamma_{ck} - U_k, 0\} dF_c(w) \quad (2.5)$$

$$rE_k(w) = w + \xi_k + \delta(U_k - E_k(w)) + \sum_c \phi_{c,k}^E \int_{\underline{w}}^{\bar{w}} \max \{E_c(w') - \gamma_{ck} - E_k(w), 0\} dF_c(w') \quad (2.6)$$

Workers' mobility decisions are characterized by a set of mobility-compatible indifference

wages, or the lowest wage that they would require to accept a job in another market. For a worker employed in k at a job that pays w , let the mobility-compatible indifference wage for market c be $w_{ck}^*(w)$, which is defined as the wage in c that gives the same expected lifetime utility, net of moving costs, as the worker's current job in k . We must also characterize reservation wages for unemployed workers. Let \underline{w}_{ck}^* be the lowest wage that an unemployed worker from market k will accept in market c . Using the value functions, reservation wages and mobility-compatible indifference wages are defined implicitly as follows:

$$U_k = E_c(\underline{w}_{ck}^*) - \gamma_{ck} \quad (2.7)$$

$$E_k(w) = E_c(w_{ck}^*(w)) - \gamma_{ck} \quad (2.8)$$

Indifference wages depend on moving costs and the relative value of the expected future offer distribution in each market. A job that is in a market that is more remote or that has a weaker local job ladder has a lower expected future value of employment to workers, so a higher wage will be needed to induce workers to move there. Conversely, workers may be willing to accept a low wage in order to move to a market that has better outside options.

For offers from one's local market, there are no moving costs and the expected future offer distribution does not change, so $w_{kk}^*(w) = w$. Using these expressions, we can rewrite the value functions above eliminating the maximization problem:

$$rU_k = b + \xi_k + \sum_c \phi_{c,k}^U \int_{\underline{w}_{ck}^*}^{\bar{w}} (E_c(w) - \gamma_{ck} - U_k) dF_c(w) \quad (2.9)$$

$$rE_k(w) = w + \xi_k + \delta(U_k - E_k(w)) + \sum_c \phi_{c,k}^E \int_{w_{ck}^*(w)}^{\bar{w}} (E_c(w') - \gamma_{ck} - E_k(w)) dF_c(w') \quad (2.10)$$

We can also derive a convenient expression for the surplus of a job to the worker, $E_k(w) - U_k$. First differentiate $E_k(w)$ with respect to w :

$$E'_k(w) = \frac{1}{r + \delta + \sum_c \phi_{ck}^E (1 - F_c(w_{ck}^*(w)))} \quad (2.11)$$

Next, I integrate using the terminal condition $E_k(\underline{w}_{ck}^*) = U_k$ to find the surplus of employment in market k

$$\Xi_k(w) \equiv E_k(w) - U_k = \int_{\underline{w}_{ck}^*}^w \frac{1}{r + \delta + \sum_c \phi_{ck}^E (1 - F_c(w_{ck}^*(\tilde{w})))} d\tilde{w} \quad (2.12)$$

Similarly, we can use the definition of indifference wages to evaluate the job-switching surplus, $\Xi_{ck}(w', w)$, or the gain from accepting a job offer that pays w' in market c . By the definition of indifference wages, $E_k(w) = E_c(w_{ck}^*(w)) - \gamma_{ck}$. Using this as a terminal condition for the integral of $E'_c(w)$ yields the job switching surplus

$$\Xi_{ck}(w', w) \equiv E_c(w') - \gamma_{ck} - E_k(w) = \int_{w_{ck}^*(w)}^{w'} \frac{1}{r + \delta + (\xi_c - \xi_k) + \sum_l \phi_{lc}^E [1 - F_l(w_{lc}^*(\tilde{w}))]} d\tilde{w} \quad (2.13)$$

Substituting these expressions into the value functions allows them to be written non-recursively as functions of parameters, the local wage distributions, and the mobility-compatible indifference wages.

$$rU_k = b + \xi_k + \sum_c \phi_{c,k}^U \int_{\underline{w}_{ck}^*}^{\bar{w}} \int_{w_{ck}^*(w)}^{w'} \frac{1}{r + \delta + (\xi_c - \xi_k) + \sum_l \phi_{lc}^E [1 - F_l(w_{lc}^*(\tilde{w}))]} d\tilde{w} dF_c(w') \quad (2.14)$$

$$\begin{aligned} rE_k(w) = & w + \xi_k - \delta \left(\int_{\underline{w}_{ck}^*}^w \frac{1}{r + \delta + \sum_c \phi_{ck}^E (1 - F_c(w_{ck}^*(\tilde{w})))} d\tilde{w} \right) \\ & + \sum_c \phi_{c,k} \int_{w_{ck}^*(w)}^{\bar{w}} \int_{w_{ck}^*(w)}^{w'} \frac{1}{r + \delta + (\xi_c - \xi_k) + \sum_l \phi_{lc}^E [1 - F_l(w_{lc}^*(\tilde{w}))]} d\tilde{w} dF_c(w') \end{aligned} \quad (2.15)$$

Using these and the definition of mobility-compatible indifference wages, equation (2.8), we can derive recursive expressions for the mobility-compatible indifference wage functions and reservation wages:

$$w_{ck}^*(w) = w + r\gamma_{ck} - r(\xi_c - \xi_k) + \delta [\Xi_c(w_{ck}^*(w)) - \Xi_k(w)] \\ + \sum_l \left[\phi_{lk}^E \int_{w_{lk}^*(w)}^{\bar{w}} \Xi_{lk}(w', w) dF_l(w') - \phi_{lc}^E \int_{w_{lc}^*(w_{ck}^*(w))}^{\bar{w}} \Xi_{lc}(w', w_{ck}^*(w)) dF_l(w') \right] \quad (2.16)$$

$$\underline{w}_{ck}^* = b + r\gamma_{ck} + \delta \Xi_c(\underline{w}_{ck}^*) \quad (2.17)$$

$$+ \sum_l \phi_{lk}^U \int_{\underline{w}_{lk}^*}^{\bar{w}} \Xi_{lk}(w', \underline{w}_{lk}^*) dF_l(w') - \sum_l \phi_{lc}^E \int_{w_{lc}^*(\underline{w}_{ck}^*)}^{\bar{w}} \Xi_{lc}(w', \underline{w}_{ck}^*) dF_l(w') \quad (2.18)$$

Equations (2.16) and (2.18) define a recursive system of $2N \times 2N$ functional equations over the $2N \times 2N$ unknown functions $w_{ck}^*(w)$ and \underline{w}_{ck}^* . By the Banach fixed point theorem, this system is a contraction (Schmutz and Sidibé, 2018), so iterating from an initial guess for the indifference wage functions will solve the system.

For intuition, suppose that there are no geographic frictions and that all markets are identical. This is equivalent to a standard on-the-job search model, and equation (2.16) collapses to $w_{ck}^*(w) = w$. Similarly, suppose that we relax search frictions but preserve the geographic structure of the model. As offer rates go to infinity, the surplus terms in (2.16) all converge to the same constant and cancel out, and we are left with $w_{ck}^*(w) = w + r\gamma_{ck}$. This is equivalent to a static spatial equilibrium model. With moving costs, bilateral search frictions, and heterogeneous cities, indifference wages take into account the fixed costs of moving and the difference in the value of expected future job offers between the origin and destination.

Labor Mobility: Job-to-job Quits

Job-to-job flows occur when a worker receives a job offer that is greater than or equal to their mobility-compatible indifference wage. A worker in city k in a job with wage w will match with a vacancy in city c at rate ϕ_{ck}^E and the vacancy will have a wage high enough for the worker to accept the job with probability $1 - F_c(w_{ck}^*(w))$. This means that the worker will accept a job offer in city c with probability

$$\lambda_{ck}^E(w) = (1 - F_c(w_{ck}^*(w))) \quad (2.19)$$

and will have a total job-to-job quit rate equal to the sum of the bilateral job-to-job quit rates, or $\sum_c \phi_{ck}^E \lambda_{ck}^E(w)$.

Labor Mobility: Job-to-Job Hires

In addition to the quit rate, the job-to-job hire rate will be necessary to evaluate firms' vacancy posting decisions. A vacancy in market c that pays a wage w will encounter a worker in market k at rate ψ_{ck} . The worker will accept the offer if their mobility-compatible indifference wage is less than w . This means that the probability of a match resulting in a hire depends on the wage distribution of employment, which will generally differ from the wage distribution of vacancies due to varying acceptance rates and job-to-job quit rates at different wages. Let $G_k(w)$ be the wage distribution of employment in market k . The probability that an employed worker in k accepts the offer is $G_k(w_{ck}^{*-1}(w))$, where $w_{ck}^{*-1}(w)$ is the inverse of the mobility-compatible indifference wage function, defined by

$$E_k(w_{ck}^{*-1}(w)) = E_c(w) - \gamma_{ck} \quad (2.20)$$

In other words, $w_{ck}^{*-1}(w)$ gives the highest wage in k that workers can make and still be willing to accept a job that pays w in c . The total hiring rate for a vacancy also depends on the acceptance decisions of unemployed workers. Unemployed workers will accept the offer if it is greater than their mobility-compatible reservation wage $w_{ck}^*(b)$. The bilateral

acceptance rate for a firm in c that pays w is the weighted sum of the acceptance rates of the employed and unemployed, or

$$\lambda_{ck}^F(w) = \frac{n_k^U}{s_k} \mathbb{1}[w] > w_{ck}^*(b)] + \frac{n_k^E}{s_k} G_k(w_{ck}^{*-1}(w)) \quad (2.21)$$

The total hiring rate of a vacancy in c that pays w is then the sum of the bilateral hiring rates, or $\sum_k \psi_{ck} \lambda_{ck}^F(w)$

Labor Mobility: Moves to Unemployment

In addition to job-to-job flows, workers have the option to pay moving costs and move to unemployment in any market. In spatial equilibrium, the unemployed will move until there are no marginal incentives to do so. As the unemployed move to more attractive markets, the offer rate in that market declines, which reduces the marginal incentive to move for other workers. This implies that the unemployed will move until the following condition holds for each city k :

$$U_k \geq \max_c \{U_c - \gamma_{ck}\} \quad (2.22)$$

Because the value of employment is weakly greater than unemployment for any offer that workers will accept, (2.22) also implies that there are no marginal incentives to move for the employed, so the only mobility for employed workers will be through job-to-job flows.

2.4.4 Firms

There is a continuum of ex-ante identical, single-worker firms that decide whether not to enter the market by posting a vacancy. Let κ be the cost of posting a vacancy. After a firm and a worker match, the productivity of the match is revealed to the firm and the worker, the worker decides whether or not to accept the job offer, and upon acceptance the firm and worker bargain over wages.

The value of posting a vacancy V_c depends on vacancy posting costs κ and the expected value of a job if a match is made. The vacancy will match with workers in market k at rate ψ_{ck} . After matching, the productivity of the job z is drawn from an exogenous local distribution $H_c(z)$. Wages will be set by bargaining, but the worker and the firm know the equilibrium outcome of the bargaining game $w_c(z)$ in advance.

If the match is with an employed worker, with probability $\lambda_{ck}^F(w_c(z))$, the expected wage will be higher than the worker's mobility-compatible indifference wage, so the worker will accept the offer (and moves to c if necessary). If the match is with an unemployed worker, then the worker accepts if the wage is higher than $w_{kc}^*(b)$. If the worker accepts the job, the firm values it at $J_c(w_c(z)|z)$, again substituting the wage bargaining solution, where:

$$rV_c = -\kappa + \sum_k \psi_{ck} \int_{\underline{z}}^{\bar{z}} \lambda_{ck}^F(w_c(z)) [J_c(w_c(z)|z) - V_c] dH_c(z) \quad (2.23)$$

$$rJ_c(w|z) = z - w - \left(\delta + \sum_k \phi_{kc}^E \lambda_{kc}^E(w) \right) J_c(w|z) \quad (2.24)$$

The value of the job to the firm $J_c(w|z)$ depends on the flow of revenue $z - w$ as well as the expectation of the value of future employment states. Jobs are exogenously ended with probability δ , and firms also take into account expected job-to-job quits. The firm's workers will receive outside offers from market k with probability ϕ_{kc}^E , and the firm knows that they will accept those offers with probability $\lambda_{kc}^E(w)$.

There is free entry, so firms will enter until there is no marginal incentive to do so. In equilibrium the following vacancy posting condition will determine the number of vacancies in each market.

$$n_c^v \equiv 0 = -\kappa + \sum_k \psi_{ck} \int_{\underline{z}}^{\bar{z}} \lambda_{ck}^F(w_c(z)) J_c(w_c(z)|z) dH_c(z) \quad (2.25)$$

In addition, note that solving (2.24) for $J_c(w|z)$ yields an expression that will be convenient in specifying the wage bargaining solution.

$$J_c(w|z) = \frac{z - w}{r + \delta + \sum_k \phi_{kc}^E \lambda_{kc}^E(w)} = (z - w) E'_c(w) \quad (2.26)$$

2.4.5 Wage Bargaining

Wages are set by bargaining between the firm and the worker. I apply the bargaining solution from [Shimer \(2006\)](#), which shows that in models of on-the-job search, standard solutions such as Nash bargaining or surplus splitting cannot be applied because the set of feasible allocations of surplus between firms and workers can be non-convex. The non-convexity arises because the expected duration of employment depends on the wage that the firm offers: if a firm raises the wage, then it is less likely that a worker will receive a better job offer in the future, so expected duration rises. Including renegotiation of wages when workers receive an outside offer, as in [Cahuc et al. \(2006\)](#), can alleviate this problem by fixing expected employment duration, but would require tracking the distribution of workers' outside offers. [Shimer \(2006\)](#) uses a particular bargaining environment that has a solution even when the payoff space is non-convex.

When a worker receives a job offer from a firm, they observe its location and productivity. Before bargaining, they first decide whether or not to accept the job, taking the expected outcome of bargaining and any moving costs they would incur by accepting the offer into account. If they accept, the worker then moves to the location of the offer before wage bargaining takes place, so during bargaining any moving costs are treated as sunk. The outside option for all workers is the value of unemployment in the location of the job, even for workers that are poached from another firm. Workers cannot leverage outside offers into raises from their current employer.

After matching, the acceptance decision, and moving, workers and firms engage in alternating offer bargaining in a game that resolves before the end of the period. There is an exogenous probability ι that bargaining breaks down at each stage for which the offer is not accepted. If bargaining breaks down, the worker becomes unemployed. The bargaining solution applies at the limit when ι approaches zero, so I assume that there are no actual breakdowns. With heterogeneous firms, [Shimer \(2006\)](#) shows that there exists an equilibrium where more productive firms offer higher wages. This equilibrium is not necessarily unique, but I will restrict attention to it.

Given this setup, [Shimer \(2006\)](#) shows the existence of an equilibrium wage function $\omega(k, z)$ that is a local maximizer of the joint surplus of workers and firms:

$$\omega(k, z) = \operatorname{argmax}_{w \in B_\varepsilon(\omega)} (E_k(w) - U_k) J_k(w|z) \quad (2.27)$$

I will follow Shimer closely in deriving the equilibrium wage. In each stage of the bargaining game, the firm will offer a wage ω^f or the worker will offer a wage ω^w . If the worker rejects the firm's offer, then with probability $1 - \iota$, the bargaining game moves on to the next stage and the worker makes the offer. The offers are part of a subgame perfect equilibrium if the firm is indifferent about accepting w^w and the worker is indifferent about accepting w^f . To define these equilibrium conditions, let $E_k(\omega)$ be the worker's value function as above but not taking wages as given. Similarly, let $J_k(\omega|z)$ be the firm's value function with variable wages. The indifference condition from the bargaining game gives us that

$$E_k(\omega^f) = (1 - \iota)E_k(\omega^w) + \iota U_k \quad (2.28)$$

$$J_k(\omega^w|z) = (1 - \iota)J_k(\omega^f|z) \quad (2.29)$$

When ι is close to zero, ω^w and ω^f converge to the wage function $\omega(k, z)$. To characterize this solution, first differentiate the indifference conditions with respect to ι :

$$\begin{aligned}
E'_k(\omega^f) \frac{\partial \omega^f}{\partial \iota} &= -E_k(\omega^w) + U_k + (1 - \iota) E'_k(\omega^w) \frac{\partial \omega^w}{\partial \iota} \\
J'_k(\omega^w|z) \frac{\partial \omega^w}{\partial \iota} &= -J_k(\omega^f|z) + (1 - \iota) J'_k(\omega^f|z) \frac{\partial \omega^f}{\partial \iota}
\end{aligned}$$

As ι converges to zero, ω^f and ω^w converge to a single wage $w_k(z)$. The conditions above then collapse to

$$\left(\frac{\partial \omega^f}{\partial \iota} - \frac{\partial \omega^w}{\partial \iota} \right) \Big|_{\iota \rightarrow 0} = \frac{-J_k(w_k(z)|z)}{J'_k(w_k(z)|z)} = \frac{E_k(w_k(z)) - U_k}{E'_k(w_k(z))}$$

Which gives the condition that equilibrium wages must satisfy:

$$\frac{J_k(w_k(z)|z)}{J'_k(w_k(z)|z)} + \frac{E_k(w_k(z)) - U_k}{E'_k(w_k(z))} = 0 \quad (2.30)$$

In other words, for the joint surplus $(E_k(w) - U_k)J_k(w|z)$ to be maximized, the elasticity of the value of the job with respect to wages must be equal in magnitude (and opposite in sign) for firms and workers.

Using equations (2.11), (2.12), and (2.26), we can write a condition that equilibrium wages satisfy

$$\frac{-(z - w)}{r + \delta + \sum_c \phi_{ck} \lambda_{ck}^E(w) + (z - w) \sum_c \phi_{ck}^E \frac{\partial \lambda_{ck}^E(w)}{\partial w}} + \int_b^w \frac{1}{r + \delta + \sum_c \phi_{ck} \lambda_{ck}^E(\tilde{w})} d\tilde{w} = 0 \quad (2.31)$$

Note that this expression is pinned down by the equilibrium wage distribution, the offer arrival rates, and the set of mobility-compatible indifference wage functions. Given a wage

distribution and an offer arrival rate, the wage bargaining solution can be computed once the workers' problem is solved.

2.4.6 Steady State

Transitions between employment states are determined by the local wage distributions of vacancies and jobs, $F_c(w)$ $G_c(w)$, bilateral offer rates ϕ_{ck}^U and ϕ_{ck}^E , and mobility-compatible indifference wages $w_{ck}^*(w)$. As workers move between cities, market tightness changes, which will affect workers' mobility decisions, firms' vacancy posting decisions, and wage bargaining. In the steady state, inflows and outflows from employment and unemployment as well as at each point in the wage distribution are equal for all locations. This gives a set of conditions on the laws of motion that must be satisfied.

The mass of unemployed workers in city k is n_k^u and the mass of employed workers in n_k^E . Define $n_k^E(w)$ as the number of workers employed at wage w in k , or

$$n_k^E(w) = n_k^E g_c(w) \quad (2.32)$$

Where $g_c(w)$ is the PDF of the wage distribution of jobs, $G_c(w)$. Inflows to unemployment are workers that are subject to exogenous job destruction shocks at rate δ . Outflows from unemployment occur at the rate at which unemployed workers find jobs that pay more than their reservation wages. In the steady state, total inflows into unemployment are equal to outflows, which gives

$$\sum_k \delta n_k^E = \sum_k n_k^u \sum_c \phi_{ck} [1 - F_c(w_{ck}^*(b))] \quad (2.33)$$

This condition on total inflows and outflows to unemployment holds for the entire economy, but for a given location it is possible for flows in and out of unemployment to be unequal because the unemployed can move to unemployment in different markets. Because of moves of the unemployed, the spatial equilibrium condition (2.22) is required to pin down the geographic distribution of the unemployed.

Similarly, at each point in the wage distribution, outflows must equal inflows in the steady state. Outflows from employment at wage w occur due to exogenous job destruction and job-to-job quits. Inflows are the total of hires from unemployment and hires from employment. Hires depends on the probability that a vacancy is for a job that pays wage w , $f_k(w)$, and how likely it is that a match results in a hire. For matches with unemployed workers in market c , the match only results in a hire if the wage is higher than the mobility-compatible indifference wage for an unemployed worker from that market. For matches with employed workers, $w_{kc}^{*-1}(w)$ is the highest wage that workers in market c can make and be willing to accept the offer of wage w in k , so all workers below this threshold accept offers. Setting inflows equal to outflows gives the following condition

$$n_k^e(w) \left(\delta + \sum_c \phi_{ck}^E [1 - F_c(w_{ck}^*(w))] \right) = f_k(w) \left[\sum_c n_c^u \phi_{kc}^U \mathbb{1}[w > w_{kc}^*(b)] + \sum_c \int_{\underline{w}}^{w_{kc}^{*-1}(w)} \phi_{kc}^E (1 - \delta) n_c^e(w') dw' \right] \quad (2.34)$$

2.4.7 Equilibrium

The steady-state equilibrium is defined by

1. Bilateral match rates ϕ_{ck} and ψ_{ck} such that $\phi_{ck} = y_{ck}/s_k$ and $\psi_{ck} = y_{ck}/v_c$, where y_{ck} is the number of bilateral matches produced by the spatial matching function given the number of vacancies in each market, v_c , and the effective mass of searchers for each market, $s_c^* = \sum_k a_{ck}s_k$.
2. A set of mobility-compatible indifference wage functions $\{w_{ck}^*(w)\}$ such that given offer rates $\{\phi_{ck}\}$ and local wage distributions $\{F_c(w)\}$, the indifference condition $E_k(w) = E_c(w_{ck}^*(w)) - \gamma_{ck}$ holds for all c and k .
3. A wage bargaining solution $w_c(z)$ such that given bilateral match rates ϕ_{ck} and ψ_{ck} , mobility-compatible indifference wages, and an employment distribution, equation (2.31) holds and the implied wage distribution, given the exogenous productivity

distribution $H_c(z)$, is equal to $F_c(w)$.

4. An equilibrium number of vacancies in each market such that given the wage bargaining solution, bilateral match rates, and mobility-compatible indifference wages, the vacancy posting condition (2.25) holds.
5. A mass of employed workers $\{n_k^e\}$, unemployed workers $\{n_k^u\}$, and wage distribution of jobs $\{G_c(w)\}$ such that inflows and outflows across employment states are equal.

Algorithm

To find the steady state of the model:

1. Guess the number of vacancies $\{v_k\}^0$ and the wage distribution in each market $\{F_c(k)\}^0$
2. Find the steady state number of employed and unemployed workers in each market, $\{n_k^E\}$ and $\{n_k^u\}$
 - Guess $\{n_k^e\}^0$ and $\{n_k^u\}^0$
 - Use the spatial matching function to compute $\{\phi_{ck}\}$ and $\{\psi_{ck}\}$
 - Given $\{F_k(w)\}$, $\{\phi_{ck}\}$, $\{\psi_{ck}\}$, the worker's problem can be solved for the set of mobility-compatible indifference wage functions $\{w_{ck}^*(w)\}$ and the value functions $\{E_k(w)\}$ and $\{U_k\}$
 - Compute the steady state values of $\{n_k^e\}$, $\{n_k^u\}$, and $\{n_k^e(w)\}$ that are consistent with the transition rates implied by the offer rates and mobility-compatible indifference wages using equations (2.33) and (2.34)
 - Check that the spatial equilibrium condition (2.22) holds. If it does not hold, reallocate unemployed workers (updating market tightness) until it does.
 - If $\{n_k^e\}^0 = \{n_k^e\}^1$ and $\{n_k^u\}^0 = \{n_k^u\}^1$, stop. Otherwise, use $\{n_k^e\}^1$ and $\{n_k^u\}^1$ as the new guess. Repeat until convergence.

3. Apply the wage bargaining solution, equation (2.31), to find the wage functions $\{w_k(z)\}$. Use the productivity distribution $\{G_c(z)\}$ to compute the wage distribution $\{F_c(z)\}^1$ that is consistent with the bargaining solution
4. Apply the vacancy posting condition, equation (2.25) to find the total number of vacancies that firms will post, $\{v_k\}^1$
5. If $\{v_k\}^0 = \{v_k\}^1$ and $\{F_k(w)\}^0 = \{F_k(w)\}^1$, stop. Otherwise, use $\{v_k\}^1$ and $\{F_k(w)\}^1$ as the new guess. Repeat until convergence.

2.5 Results

In this section I discuss results from a simple version of the model with two symmetric cities. This setup loses several interesting features of the model, such as geographic heterogeneity or the diversion effect resulting from connections to third cities. However, this version greatly simplifies the parameter space and reduces the dimensionality of the decision rules in the model while still allowing me to study the effect of non-local job opportunities on labor market outcomes.

I calibrate the simple version of the model to match several aggregate moments regarding within-metro and across-metro job-to-job flows. This calibration is of interest itself because it provides an approximation of the average frictional connections between US labor markets. While the current implementation is very preliminary, I do find that moving costs required to explain observed mobility are far lower than estimates in the literature, consistent with [Schmutz and Sidibé \(2018\)](#). To explore this result further, I use the expression for the workers' mobility-compatible indifference wages to show that the present value calculations required to translate the gains from moving into estimates of moving costs are fundamentally different once search frictions are taken into account.

2.5.1 Environment and Parameters

I study an economy that has two cities with identical local productivity distributions and symmetric frictions. I normalize within-market search effectiveness a_{cc} to be equal to one. In a symmetric model population is equal in equilibrium, so the size-dependent amenity ξ_k

is constant and can be normalized to zero. With these simplifications, only two parameters are needed to characterize the geographic frictions in the economy, namely, cross-market search effectiveness $a_{12} = a_{21}$ and moving costs $\gamma_{12} = \gamma_{21}$.⁷

I assume that the productivity of vacancies is drawn from a Beta distribution. The Beta distribution is very flexible, yet requires few parameters, which will facilitate calibration to data on local wage distributions in the future (Meghir et al., 2015; Schmutz and Sidibé, 2018). Vacancies in each city have an identical Beta distribution over the productivity support $[b, \bar{z}]$ with parameters α and β . The PDF is

$$h(z) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathbf{B}(\alpha, \beta)}, \quad x = \frac{z-b}{\bar{z}-b}$$

where $\mathbf{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, and Γ is the Gamma function. I discretize the distribution over a support with 36 linearly spaced points. At each point in the support, I use the exact CDF given by the Beta distribution.

I assume total population is two. Because the cities are symmetric, they will have equal population in equilibrium, though this will not be the case more broadly with asymmetric cities. To compute the matching probabilities, I assume that the matching function has a Cobb-Douglas form

$$m(v_c, s_c^*) = v_c^\theta s_c^{*1-\theta} \tag{2.35}$$

The parameterization of the model is detailed in Table 2.6. I calibrate the majority of parameters to match model moments detailed in Table 2.7. A small number of parameters are fixed ex-ante, namely the discount rate r , which is set to the equivalent of a 5% annual discount rate, and the job destruction rate δ , which I peg to the average quarterly layoff rate

⁷In the more general case of N symmetric cities, there are $(N^2 - N)/2$ moving cost parameters and the same number of bilateral search effectiveness parameters. This can be reduced further by parameterizing the frictions with respect to observable variables such as physical distance as in Schmutz and Sidibé (2018)

from the Job Opening and Labor Turnover (JOLTS) in 2015. In addition, θ , which is the elasticity for vacancies in the matching function, is fixed. I find that in general equilibrium the effect of θ on the model solution are very similar to changes in the vacancy posting cost, which can cause issues with convergence in the calibration, so I simply set θ to .5.

For calibrated parameters, I target moments from the J2J and other data sources for the year 2015. For moving costs and bilateral search effectiveness, I target the ratio of the change in average earnings within and across metro areas and the ratio of job-to-job flows across metro areas. In the J2JOD data, job switches across metros are associated with a 40% larger gain in average earnings, and the ratio of job-to-job flows across metro areas to those within metro areas is roughly 63%. I target several other moments from the J2J data including the average quarterly rate as a percentage of employment, and the share of hires from employment in order to pin down the relative search intensity of employed workers, α , and matching efficiency A . To pin down the unemployment benefit, I target the average quarterly U3 unemployment rate in 2015, 4.99%. For the productivity distribution, I target the mean and standard deviation of average quarterly earnings for full-time workers in 2015 using data from the Current Population Survey. I normalize productivity and earnings so that the maximum level of productivity \bar{z} in the model, which I set to 10, is equivalent to the CPS topcode for reported earnings, which is \$150,000 annually or a quarterly earnings of \$37,500. Lastly, I target an average markdown of wages from productivity of .8 in order to pin down the vacancy posting cost κ , which is in line with estimates of average markdowns from [Berger et al. \(2022\)](#).

Moments in the calibrated model and the data are listed in [Table 2.7](#). While most model moments generally close to their targets, I am less able to match the quarterly hire rate and the share of hires accounted for by J2J flows ⁸. There are tradeoffs in meeting all of the targeted moments simultaneously. For example, lowering the search intensity of employed workers could reduce the share of hires from employment, but would also reduce the mean and variance of wages and raise unemployment by increasing reservation wages. While

⁸The difference between the share of hires from employment in the model and data may be overstated here. To calculate this moment in the data I use all new hires from the QWI as the denominator. In the context of job-to-job flows, new hires into main jobs are commonly used instead of all new hires, for example in [Haltiwanger et al. \(2018\)](#). Using this definition gives a share of hires from employment that is roughly 40% in the data.

Table 2.6: Parameters

Parameter	Value	Description	Source/Target
Fixed:			
r	.0129	Interest Rate	5% annual discount rate
θ	.5	Matching Function Elasticity	-
δ	.0389	Job Destruction Rate	3.89% Quarterly Layoff Rate
Calibrated:			
A	.4205	Matching Efficiency	Quarterly Hire Rate
σ	.2615	Search Intensity of Employed	J2J Share of Hires
κ	1.615 (\$6,105)	Vacancy Posting Cost	Average Markdown
b	.820 (\$3,075)	Unemployment Benefit	Unemployment Rate
α	6.663	Productivity distribution (Beta)	Average Wage
β	18.965	Productivity distribution (Beta)	Standard Deviation of Wages
\bar{z}	10 (\$37,500)	Maximum productivity	-
γ_{ck}	2.4281 (\$9,105)	Moving Costs	Relative job-switching gain
a_{ck}	.8566	Bilateral search effectiveness	Cross Market J2J share

in absolute terms there are fewer hires and a greater share of hires from employment in the model than observed empirically, the calibration matches the share of nonlocal J2J flows quite closely. Similarly, the wage gain from nonlocal J2J flows is close to that in the data, but somewhat lower, suggesting that moving costs may be underestimated in this calibration.

One notable result of the calibration is that the geographic frictions are relatively weak. Moving costs are about 68% of average quarterly earnings (roughly \$9,000) and search across markets is about 86% as effective as search within markets. This estimate of bilateral search effectiveness is consistent with [Schmutz and Sidibé \(2018\)](#), who find that for the median French city, the total offer rate from non-local cities is about 90% of the local offer rate. While the estimate of moving costs is quite low here, I do expect moving costs to be much less than consensus estimates in the literature, for reasons which I discuss this in the next section.

2.5.2 Mobility-Compatible Indifference Wages

Figure 2.2 shows the equilibrium mobility-compatible indifference wage function, expressed as premium over current wages. Because the locations are symmetric, the wage

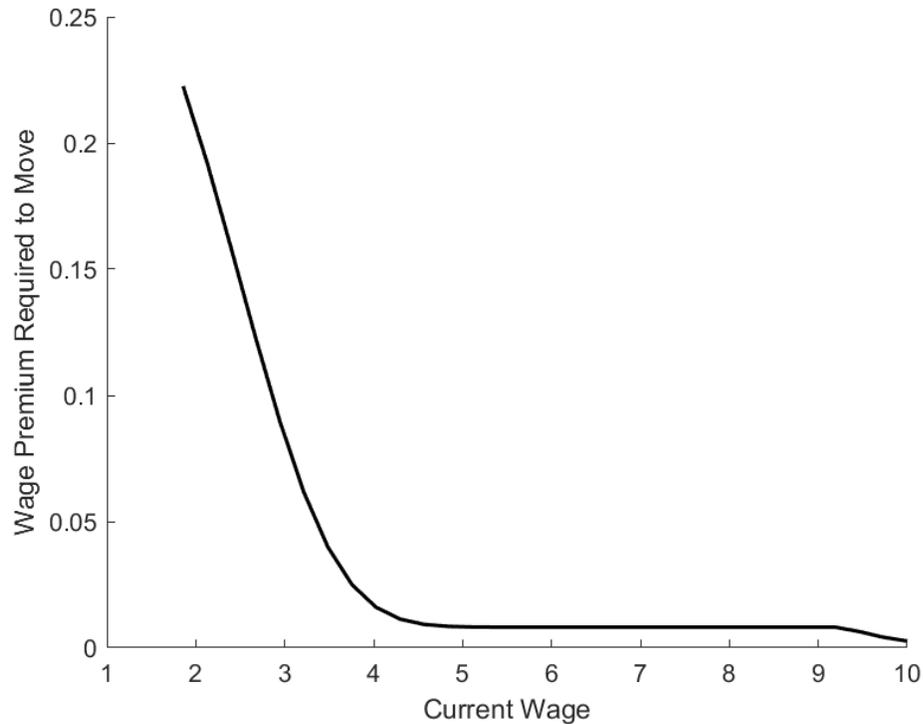
Table 2.7: Calibration: Targeted Model Moments

Moment	Model	Data	Source
Unemployment Rate	4.99%	5.3%	Average U3 (BLS), 2015
Hires/Employment	9.07%	14.7%	J2J statistics, 2015
J2J Share of Hires	.548	.268	J2J statistics, 2015
J2J Across / J2J Within	.6236	.6323	J2J statistics, 2015
Wage Gain Across / Wage Gain Within	1.361	1.462	J2J statistics, 2015
Average Quarterly Earnings	\$13,412	\$13,335	CPS-IPUMS, 2015
S.D Quarterly Earnings	\$4,975	\$4,899	CPS-IPUMS, 2015
Average Markdown	.818	.80	-

premium depends only on moving costs and the workers' current position in the wage distribution. The wage premium workers require to move is decreasing in the current wage. As wages reach the upper limit of the wage distribution, the wage premium asymptotes to $\gamma_{ck} * r$, or the full amortization of moving costs. Intuitively, in order to accept an offer from the other market, the worker needs to be compensated for moving costs. However, the opportunity cost of moving is not simply the worker's current job, but the value of their expected future employment, which includes better local job offers that workers may encounter in the future and which they can accept without incurring moving costs. Near the top of the wage distribution, the probability of receiving a local offer that dominates the worker's current job is vanishingly small, so moving costs are fully amortized. Workers lower in the job ladder encounter better local offers at a higher rate, so they amortize moving costs over a shorter effective horizon.

One immediate lesson from Figure 2.2 is that the average gain from moving is not sufficient to identify moving costs. Workers at different positions in the job ladder amortize moving costs over different effective horizons, so the discounted present value of the full amortization only applies for workers near the top of the job ladder. This suggests that previous work estimating moving costs using spatial dispersion in wages or the gains from moving such as [Kennan and Walker \(2011\)](#), [Artuç and McLaren \(2015\)](#), or [Ransom \(2018\)](#) are likely biased upwards, because the present value calculation does not take into account the reduced effective horizon due to the prospect of encountering a better job locally. This is precisely what drives the findings of [Schmutz and Sidibé \(2018\)](#), who show that accounting for on-the-job search results in estimated moving costs that are an order of magnitude

Figure 2.2: Mobility-Compatible Indifference Wage Premium



The mobility-compatible indifference wage premium is the difference between the current wage and the lowest wage in the other market that a worker is willing to accept.

lower than typical estimates in the literature. To further support this point, in the calibrated model the average quarterly wage gain for nonlocal job switches is roughly 3,580 greater than that for local job switches. At a 1.29% quarterly discount rate, the expected value of this wage increase over a 25-year horizon is \$200,803. For an infinitely lived agent, the expected value asymptotes at \$286,440. These figures imply an upper range of moving costs that an agent in a static labor market would require in order to be indifferent between switching jobs locally and nonlocally that is in line with the above estimates from the literature based on static expected value calculations. The dynamic features of the model are thus key in enabling much lower moving costs to limit observed geographic mobility rates.

The declining relationship between current wages and the mobility-compatible indifference wage premium also suggests that the local labor market is a relatively more important source of opportunities for workers that are lower in the job ladder. To explore this further,

Figure 2.3: Relative Labor Mobility along the Job Ladder

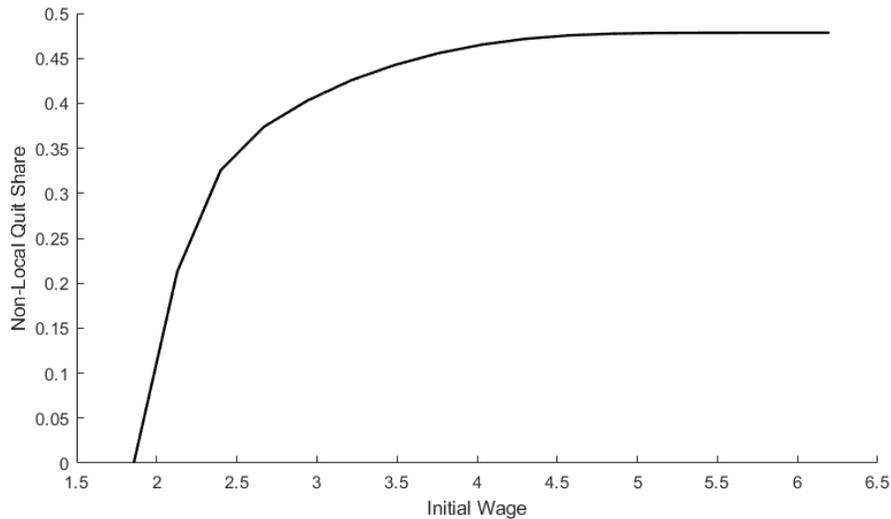


Figure plots the share of job switches that are across cities at each point in the initial wage distribution.

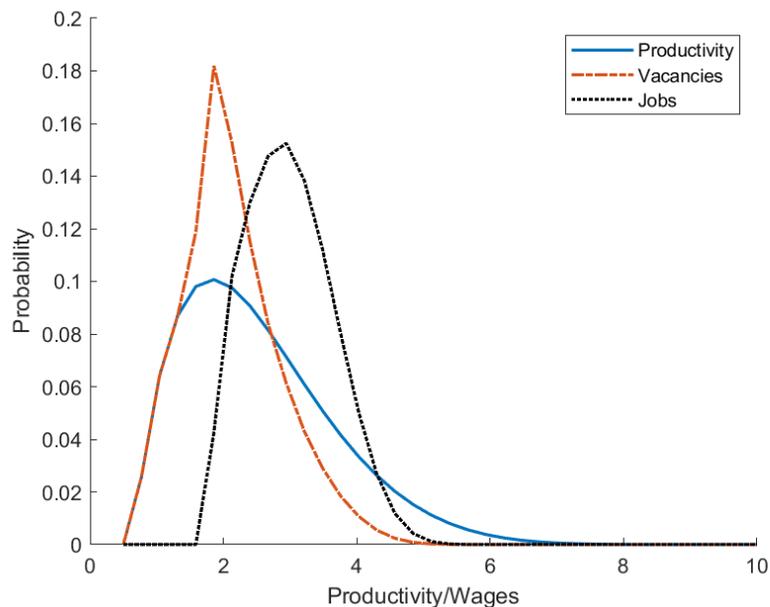
Figure (2.3) plots the share of job switches that take place across cities at each point in the initial wage distribution. At the low end of the wage distribution, almost all job-to-job flows are local. As initial wages rise, the non-local quit share approaches (though never quite reaches) 50%.

2.5.3 Wages and Markdowns

One of the advantages of the general equilibrium model I develop is that the posted wages are endogenous. In addition to increasing offer rates and speeding transitions up the job ladder, relaxing geographic frictions can affect firm rents.

Figure (2.4) plots several probability density functions from the benchmark model. First is the exogenous density of the productivity of vacancies $h_c(z)$. After matching and realizing the productivity of the match, the firm and worker bargain over wages. Combining the productivity density and the bargaining solution gives the wage density of vacancies, $f_c(z)$. The wage density has greater mass at the lower end than the productivity density due to markdowns, which I plot in Figure 2.5. The two densities coincide below workers' reservation wages, but for any job that workers are willing to accept, the firm extracts rents which are increasing in the productivity of the match. Because workers search on-the-job,

Figure 2.4: Distribution of Productivity and Wages



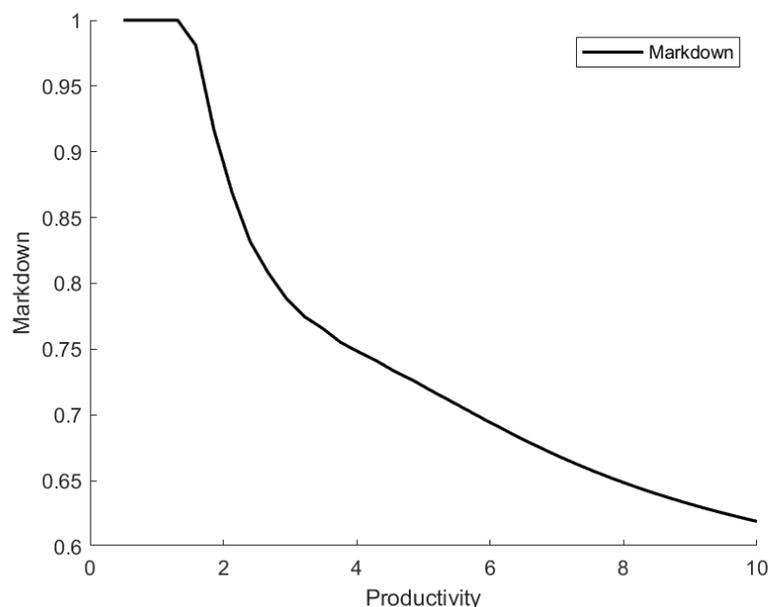
the wage distribution of jobs is shifted back to the right relative to the wage distribution of vacancies.

In interpreting Figure 2.4, recall that productivity and wages are normalized so that the maximum value, 10, is equivalent to quarterly earnings of \$37,500, and the average wage, 3.57, is equivalent to \$13,412. One weakness apparent in the current calibration is that there is very little skewness in the wage distribution of employment, while wages in the data are characterized by a long right tail. For tractability I omit worker heterogeneity and assume that all workers in the same location bargain with the same outside option, and both of these features compress the wage distribution.

2.5.4 Counterfactuals

In the benchmark model, I calibrate geographic frictions match aggregate labor market moments. To study the contribution of the frictional connections between cities to labor market outcomes, I compare the benchmark case to two counterfactuals. In the first case (closed), I assume that labor markets are strictly local: workers can only encounter jobs locally and there are no moves of the unemployed. For the closed case, I also report partial

Figure 2.5: Markdown of Wages from Productivity



equilibrium moments in which I fix vacancies to their benchmark level. In the second case (open), I assume that there are no frictions impeding labor mobility by reducing moving costs to zero and setting bilateral search effectiveness across markets to one.

Table 2.8 reports comparative statics between the benchmark, closed, and open economies. In the economy with no labor mobility, average wages are roughly 7% lower and the unemployment rate is 30% higher. The difference in wages is related both to the strength of the job ladder and the change in workers' bargaining positions: in the closed economy there are fewer productivity enhancing job switches, which reduces aggregate productivity. Workers' outside options are also weaker, which reduces wages conditional on the productivity of the job and leads to lower average markdowns. Removing geographic frictions entirely similarly results in an increase in average wages by about 6% and lower unemployment by about 18%. The effects of removing geographic frictions are smaller as the calibrated geographic frictions are relatively weak.

To show the importance of the response of labor demand to changes in geographic frictions, I also report partial equilibrium moments for the closed economy in which I fix the level of vacancies. The reduction in wages and increase in unemployment are greatly overstated

relative to the general equilibrium case. Intuitively, restricting labor mobility reduces the probability of job-to-job quits and lowers the wages that firms pay given their productivity, which both improve the value of a given job to the firm and increase the incentives for vacancy posting. The resulting increase in labor demand mitigates the decline in wages and job switching in the general equilibrium case.

Table 2.8: Comparative Statics

	Benchmark	Closed (PE)	Closed	Open
Avg. Wage	3.576	2.693	3.336	3.784
Avg. Productivity	4.072	3.226	4.021	4.296
Avg. Markdown	0.818	0.786	0.790	0.823
Unemployment Rate	4.99%	12.66%	6.56%	4.14%
Hire Rate	9.07%	6.63%	8.12%	9.84%
J2J Rate	4.97%	3.63%	4.02%	5.74%
J2J Across / J2J Within	0.624	-	-	1
Wage Gain Within	0.218	0.210	0.244	0.271
Wage Gain Across	0.295	-	-	0.271

Lastly, I can decompose the effect of changing frictions on average wages into two channels: improving job ladders means that workers are in better jobs, and improving outside options means that workers have a better bargaining position and higher wages conditional on the job. By fixing the productivity distribution of jobs and changing wages and vice versa, I can decompose the change in wages into these two channels.

Moving to the closed economy reduces average wages by about 7%. Only about half of the change in average wages is due to weaker job ladders and lower average productivity, and the remainder - 50.34% - is due to change in markdowns. When I relax geographic frictions entirely, the change in average wages is smaller but a higher share is due to reductions in firm rents. These results suggest that changes in employer market power are a quantitatively important margin of adjustment in response to changes in geographic frictions.

It is worth noting that the economic growth induced by relaxing geographic frictions arises despite there being no role for net geographic labor reallocation: productivity is equal in both cities, and there are no economies of scale, so there is geographic misallocation that can inhibit productivity as in [Hsieh and Moretti \(2015\)](#) or [Heise and Porzio \(2019\)](#). Rather, the gains to productivity and wages are due to the increase in gross labor flows: reducing

Table 2.9: Decomposing Change in Average Wages

	Percent Change in Average Wages from Benchmark		
	Total	Δ Markdowns	Δ Productivity
Closed (Partial Equilibrium)	-9.85%	40.45	59.55
Closed	-6.72%	50.34	49.66
Open	5.81%	53.62	46.38

search frictions and moving costs between cities lowers the overall degree of labor market frictions, improving job finding rates. This reduces the level of frictional unemployment and leads to more rapid climbs up the job ladder. It is this increased rate of job-to-job flows that accounts for the increased productivity, as more workers are in better jobs. The increase in wages is greater than the increase in productivity because reducing the effective degree of search frictions improves workers' bargaining positions and limits firm rents.

2.6 Conclusion

I have shown that churning of workers across U.S. metro areas is a major feature of the labor market: there are roughly 60% as many job-to-job flows across metro areas as within metro areas. Furthermore, job-to-job flows across metro areas are largely driven by churn: the contribution of net flows is less than 10% of all cross-metro job-to-job flows. I argue that this continual reallocation of workers is driven by variation in individual workers' job opportunities, which are only partly captured by local labor market conditions. Furthermore, bilateral job-to-job flows are well characterized by a gravity model in which intercity distance is a key explanatory variable. This suggests that local labor markets' geographic positions are an important determinant of the job opportunities that workers face and how they evolve in response to changing economic conditions locally or elsewhere.

I also introduce a general equilibrium model of spatial on-the-job search in order to understand the implications of heterogeneous connections between metro areas for workers' outcomes and the transmission of local shocks across markets. The model extends existing work on the partial equilibrium implications of bilateral search frictions and moving costs by incorporating firms' vacancy posting decisions and strategic bargaining over wages. This makes the the wage distribution, and importantly, markdowns of wages from produc-

tivity, endogenous.

I calibrate a simple version of the model to match the aggregate share of job-to-job flows across metro areas and the relative gain from switching jobs across metro areas. Calibrated moving costs are much lower than typical estimates in the literature, and I provide an intuitive explanation for this from the model: the expected value calculations that relate the observed gains from moving to moving costs depend on the workers' position in the local job ladder, and the full amortization of moving costs is only relevant for workers that are near the top of the job ladder. This means that moving cost estimates that do not account for search frictions are likely overstated.

Lastly, I use the model to explore the contribution of the calibrated frictional connections between markets to wages and employment. In these exercises, cities are symmetric, so there is no role for net geographic labor reallocation to contribute to aggregate wage or productivity. Despite this, I find that relative to an economy with no labor mobility, average wages are 6% higher and unemployment is 18% lower. These gains arise from the interaction of geographic mobility and search frictions: improving matching across geographically distinct markets reduces the effective degree of search frictions in the economy, leading to lower frictional unemployment and improving job ladders, both of which contribute to higher earnings. The endogeneity of wages in the model allows me to decompose the change in average wages into two channels. First, by increasing offer rates, connections to other markets strengthen job ladders and improve the distribution of productivity of jobs. Second, connections to other markets improve workers' bargaining position and result in higher wages conditional on the productivity of the job. I find that this second channel is quantitatively significant: lower markdowns account for about half of the difference in average wages relative to an economy with no labor mobility.

While the simple version of the model has interesting aggregate implications, there is room in future work to explore heterogeneous frictional connections between U.S. labor markets. Calibration of a multi-city model involves substantial computational challenges, as the dimensionality of the model, and thus computational time and the number of parameters, increases exponentially with the number of cities. In addition, one advantage of the general equilibrium framework is that the model has explicit predictions for the evolution of

labor market flows and market tightness in response to local shocks. While the spatial environment does not permit studying a stochastic equilibrium⁹, transition dynamics between steady states are feasible and would offer a structural companion to the VAR-based analysis of previous work that estimates the importance of migration as a margin of adjustment in response to local shocks (Blanchard and Katz, 1992; Dao et al., 2017).

⁹From Moscarini and Postel-Vinay (2013), the canonical job ladder model has a dynamic equilibrium when workers' rankings of firms is the same in all aggregate states. This explicitly does not hold in a spatial environment: when aggregate states change, workers update their mobility-compatible indifference wages, which reflects their ranking of firms.

2.7 Appendix to Chapter 2

In this section I briefly outline an alternative setup where firms are ex-ante heterogeneous and the productivity distribution of vacancies is endogenously determined by their vacancy posting decisions. This adds complexity to the problem of finding the equilibrium wage distribution because the number of vacancies at each point in the productivity distribution needs to be determined, but introduces a potentially interesting margin of adjustment: the geographic sorting of firms by productivity.

2.7.1 Firms' Value Functions

There is a continuum of single-worker firms that decide whether or not to enter the market by posting a vacancy. Firms have heterogeneous productivity z drawn from a support that is common across markets, and the productivity distribution of vacancies will be determined by the posting decisions of firms. Let κ be the cost of posting a vacancy. The expected value of posting a vacancy $V_c(z)$ depends on the cost, how likely the vacancy is to yield a hire, and the value of the job to the firm. For a firm in city c with productivity z , the value of posting a vacancy is given by $V_c(z)$ and the value of a job to the firm is $J_c(w|z)$

Firms pay a fixed cost κ to open a vacancy in their market, c . A vacancy results in a match with a worker from market k at rate ψ_{ck} . The probability that a given match results in a hire and the value of the hire depend on the wage that the firm pays. Wages will be determined by bargaining, but I assume that firms and workers know the outcome of the bargaining in advance. Let $w_c(z)$ be the equilibrium wage function of a firm in market c with productivity z . The bilateral job filling rate for a firm with productivity z is then $\psi_{ck}\lambda_{ck}^F(w_c(z))$. Firms' vacancy posting decisions depend on the wages they expect to pay, so for $V_c(z)$ we can use the value of the job given the equilibrium wage function, $J_c(w_c(z)|z)$.

$$rV_c(z) = -\kappa + \sum_k \psi_{ck}\lambda_{ck}^F(w_c(z)) [J_c(w_c(z)|z) - V_c(z)] \quad (2.36)$$

$$rJ_c(w|z) = z - w - \left(\delta + \sum_k \phi_{kc}^E \lambda_{kc}^E(w) \right) J_c(w|z) \quad (2.37)$$

The value of the job to the firm $J_c(w|z)$ depends on the flow of revenue $z - w$ as well as the expectation of the value of future employment states. Jobs are exogenously ended with probability δ , and firms also take into account expected job-to-job quits. The firm's workers will receive outside offers from market k with probability ϕ_{kc} , and the firm knows that they will accept those offers with probability $\lambda_{kc}^E(w)$.

There is free entry, so firms will enter until there is no marginal incentive to do so. This gives the following vacancy posting condition that holds for each level of productivity z :

$$V_c(z) = 0 \quad \implies \quad \frac{\kappa}{\sum_k \psi_{ck} \lambda_{ck}^F(z)} = J_c(w_c(z)|z) \quad (2.38)$$

Chapter 3: Quality Adjustment at Scale: Hedonic vs. Exact Demand-Based Price Indices

With Gabriel Ehrlich, John Haltiwanger, Ron Jarmin, David Johnson, Luke Pardue, Matthew Shapiro, and Laura Yi Zhao¹

3.1 Introduction

Retail businesses create item-level data on the prices and quantities of the goods that they sell. Such data form the basis for re-engineering key economic indicators by building internally consistent aggregates of value, volume, and price directly from item-level transactions. Aggregation of transactions data could supplant traditional surveys and enumerations conducted by statistical agencies (see, e.g., [Ehrlich et al., 2022](#)). There are many potential advantages to such a re-engineering. One is to address the issue of rapid product turnover, which is likely to be associated with quality improvements. Current statistical agency procedures for measuring prices inadequately address such turnover. This paper considers

¹Ehrlich: University of Michigan; Haltiwanger: University of Maryland and NBER; Jarmin: U.S. Census Bureau; Johnson: University of Michigan; Olivares: University of Maryland; Pardue: University of Maryland; Shapiro: University of Michigan and NBER; Zhao: University of Maryland and Bank of Canada. Laura Zhao worked on this project when she was a doctoral student at the University of Maryland. This research was produced with financial support of the Alfred P. Sloan Foundation and the additional support of the Michigan Institute for Data Science, the Michigan Institute for Teaching and Research in Economics and the U.S. Census Bureau. Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. We also use NPD data housed at the U.S. Census Bureau. All results using NPD data have been reviewed to ensure that no confidential information has been disclosed (CBDRB-FY19-122, CBDRB-FY21-074, and CBDRB-FY23-067). Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the view of the U.S. Census Bureau.

scalable procedures for constructing price indices using item-level transactions data that can account for entering and exiting goods as well as changing consumer valuations of product attributes.

The use of high-frequency, item-level sales data to produce accurate inflation measures also requires incorporation of advances in index number and economic theory. We consider two complementary approaches: hedonics and demand-based models. Both approaches suggest that quality improvement is widespread across a large range of consumer goods, including in categories in which technological progress is not immediately visible.

Our preferred hedonic approach builds on the insights of [Erickson and Pakes \(2011](#), hereafter “EP”), who develop a novel method of calculating hedonic price indices that can account for changing valuations of both observable and unobservable product characteristics. High-frequency, item-level transactions data with prices, quantities, and attributes permit the implementation of the EP hedonics approach with superlative price indices (such as the Fisher or Tornqvist) in real time using internally consistent expenditure weights. We show that the EP methodology has important advantages relative to commonly used alternative hedonic methods such as the time dummy method.

Our demand-based approach builds on the exact price indices developed from theoretical models of consumer demand: the Sato-Vartia price index ([Sato, 1976](#); [Vartia, 1976](#)); the [Feenstra \(1994\)](#) adjustment to the Sato-Vartia index, which adjusts for quality change from product entry and exit (denoted the Feenstra index hereafter); and the Constant Elasticity of Substitution (CES) Unified Price Index (CUPI) developed in [Redding and Weinstein \(2020\)](#), which adjusts the Feenstra index for changing consumer preferences. The demand-based approaches have the attractive feature that they yield exact price indices under certain sets of assumptions. Moreover, in principle these methods impose sufficient structure that they can be implemented without attribute data beyond a product taxonomy. That said, a challenge for the demand-based models is that they assume a specific functional form for consumer preferences, which may or may not hold in the data.

A common feature of the frontier research methods using both hedonic and demand-based approaches is that they can account for changing consumer valuations of products or prod-

uct characteristics. In principle, the CUPI captures both quality change due to product turnover and time-varying product appeal over the course of products' time in the marketplace, without directly using detailed product attributes. The EP approach also reflects changing consumer valuations of various product attributes over time as they translate into the changing mapping between prices and characteristics.

We implement the hedonic and demand-based approaches at scale using item-level transactions data from two major sources. The first is from the NPD Group, which covers a wide range of general merchandise goods from bricks and mortar and online retailers. In this paper, we construct price indices for a select number of product groups: memory cards, headphones, coffee makers, boys' jeans, and work and occupational footwear. The NPD data include rich product attributes, which facilitate the implementation of the EP methodology. The second platform we use is the Nielsen Marketing data provided by the Kilts Center at the University of Chicago Booth School of Business, which covers a wide range of food products from grocery stores, discount stores, pharmacies and liquor stores. The Nielsen data have only sparse information on product attributes, which we address by adapting the EP methodology to a machine learning (ML) framework.

Consistent with the literature using scanner data, we find rapid product turnover, along with rich post-entry product life-cycle dynamics. Products' market shares peak several quarters after entry, while on average prices decline monotonically after entry. We also find evidence of substantial quality adjustment in price indices using either hedonic or demand-based approaches across the full range of product groups we consider. The magnitude of the quality adjustment is greater for high-tech goods such as memory cards and headphones, but we find that quality adjustment is pervasive for food product groups as well. While the latter result might be surprising, our findings are consistent with the changing and increasing variety of food products available over time.

We find that the EP method of hedonic adjustment, which can account for unobservable product characteristics, yields more systematic evidence of pervasive quality adjustment than the time dummy hedonic method. The EP method also produces a meaningful improvement in our hedonic regressions' goodness of fit. Currently, the Bureau of Labor Statistics (BLS) uses hedonic adjustment for only about 7.5 percent of goods and com-

modities in the CPI (our estimates based on [Bureau of Labor Statistics, 2023](#)). Our results extend the EP methodology from the CPI database, which does not contain item-level quantities or expenditures, to transactions data with internally consistent prices and quantities. We show that the advantages of the EP approach extend to our data environment, bolstering the case for the applicability of these data and methods in official statistics. With the techniques used in this paper and with the ubiquitous availability of item-level transactions data, there is increasingly little reason why statistical agencies should not quality-adjust most goods and services.

Among the demand-based methods, we find that the Feenstra index systematically measures lower price inflation than the Sato-Vartia. This result suggests that product turnover is associated with quality improvement. We find that the CUPI, which generalizes the Feenstra index to allow for changing product appeal over product life cycles, implies substantial additional quality adjustment relative to the Feenstra index.

A challenge in implementing the CUPI is that two of its three components are unweighted geometric means. These terms are sensitive to the inclusion of goods with very small quantities sold or market shares, which is one reason that unweighted indices are generally discouraged in the index number literature. [Redding and Weinstein \(2020\)](#) employ a reallocation procedure by which they move a subset of goods out of the CUPI's unweighted geometric mean terms and into the [Feenstra \(1994\)](#) adjustment term using what we term a *common goods rule* based on the durations of products' time in the marketplace. Applying a common goods rule brings the CUPI's measurement of price changes closer in line with other indices. Our results suggest more research is needed to provide guidance about how to define common goods rules.

3.2 Data

This section provides an overview of the two data sets that we use to compute price indices. The first comes from the NPD Group and the second comes from Nielsen. For both data infrastructures, we aggregate the item-level transactions data to the quarterly, national level and focus on quarterly price indices. This approach facilitates comparing traditional, hedonic, and demand-based price indices in a manner consistent with the recent literature.

3.2.1 NPD Data

We use proprietary data that the NPD Group provided to the U.S. Census Bureau. They consist of monthly sales and quantity data at the product-store level from 2014 through 2018.² The NPD group tracks more than 65,000 retail stores, including online retailers. The retail stores cover a wide range of general merchandise products. The NPD data analyzed here consist of five broad product groups, within which we conduct our analysis separately: memory cards, coffee makers, headphones, boys' jeans, and work/occupational footwear (hereafter simply "occupational footwear"). The NPD data have unique item-level identifiers that are consistent cross-sectionally and over time. We aggregate the item-by-store-level observations to the national product-quarter level and calculate total quantity sold and average price for each product-quarter. The item-level data cover tens of thousands of product-quarter-level observations.

The NPD data contain detailed and organized information on the characteristics of each product. Beyond basic information such as product category and brand, these characteristics include details on different types of products within the broader categories (e.g., on-ear vs. in-ear headphones; coffee vs. espresso machines) and the features or attributes of different products (e.g. built-in grinders or auto-on/off settings for coffee makers). In some cases, the attributes include continuous variables, such as memory size for memory cards. We use the detailed product characteristics in the estimation of hedonic price indices and to group products into subcategories in our estimation of nested CES models.

Table 3.1 displays average item-level product turnover rates for each product group. Each of the groups exhibits a high degree of product turnover, ranging from 4.5 percent to 13.5 percent in terms of quarterly entry and exit rates. Figure 3.1 presents life-cycle dynamics of product market shares and prices within these product groups. The illustrated statistics are mean log differences from the product-specific initial values upon entry. Prices decline steadily after entry, while market shares exhibit a hump-shaped pattern post-entry.

²Month definitions follow the National Retail Federation (NRF) calendar ([National Retail Federation, 2023](#)). The NRF calendar is a guide for retailers that ensures sales comparability between years by dividing a year into months based on a 4 weeks-5 weeks-4 weeks format. The layout of the calendar lines up holidays and ensures the same number of Saturdays and Sundays in comparable months across years. The NRF calendar thus ensures the comparability of the aggregated sales over time.

The post-entry patterns of market shares differ considerably across product groups. For example, while memory cards, coffeemakers and headphones all peak after 3 quarters, headphones decline much more rapidly than memory cards or coffeemakers. Magnitudes at the peak are large but also differ by product group. For memory cards and coffeemakers the peak is about 300 log points relative to the first quarter while for headphones the peak is about 200 log points.

Taken together, these findings highlight two important features of the data. First, there is considerable item-level product turnover that is a potentially important source of changing product quality. Second, post-entry dynamics suggest that it may be important for methods of quality adjustment to account for time-varying product appeal. Both the hedonic and demand-based approaches we consider can account for such variation.

3.2.2 Nielsen Data

We use the Nielsen Retail Scanner data (also referred to as RMS) from the Kilts Center for Marketing at the University of Chicago Booth School of Business for the period 2006 to 2015. The data consist of over 2.6 million products identified by the finest level of aggregation—12-digit universal product codes (UPCs) that uniquely identify specific goods. The Nielsen Retail Scanner data are collected from over 40,000 individual stores from approximately 90 retail chains in over 370 metropolitan statistical areas (MSAs) in the United States. Total sales in Nielsen RMS are worth over \$200 billion per year and represent 53% of all sales in grocery stores, 55% in drug stores, 32% in mass merchandisers, and 2% in convenience stores.

Nielsen organizes item-level goods into 10 departments, over 100 product groups, and over 1,000 product modules. A typical department is, for example, dry grocery, which consists of 41 product groups, such as baby food, coffee, and carbonated beverage. Within the carbonated beverage product group, there are product modules such as soft drinks and fountain beverage. We have classified the product groups into food and nonfood categories based on our own judgment in communication with researchers at the BLS. Tables [3.A.6](#) and [3.A.7](#) show our classifications into the food and nonfood categories. Appendix [3.8.1](#) describes how we clean and prepare the data for analysis at the product group-quarter level.

We focus on results for the Nielsen data’s food product groups in the main text because we estimate that the data’s coverage is more extensive and tracks economywide time trends more closely for those groups than for the nonfood product groups. Appendix 3.8.2 describes the analysis underlying that conclusion.

3.3 Conceptual Framework

Time series price indices aim to measure approximately or exactly the change in the cost of living between two or more time periods. One important challenge in constructing price indices from item-level data is the substantial pace of product turnover that we documented in the previous section. Another important challenge is that consumer preferences over products or valuations of different product characteristics may vary over time. Traditional “matched-model” price indices do not capture quality change from such product turnover or from changing relative product appeal. In contrast, the hedonic and demand-based indices we construct have the potential to capture these changes.

3.3.1 Traditional Price Indices

Our empirical work in this paper focuses on so-called geometric price indices, which are weighted averages of log price changes. We will focus in particular on the Tornqvist index, given by

$$\ln \Psi_t^{TQ} \equiv \sum_{k \in \mathbb{C}_t} \frac{s_{kt-1}^* + s_{kt}^*}{2} \ln \frac{p_{kt}}{p_{kt-1}}. \quad (3.1)$$

The set \mathbb{C}_t in equation (3.1) is the set of all “continuing” or “common” goods that are sold both in period t and in period $t-1$, while s_{kt}^* and s_{kt-1}^* denote product k ’s share of expenditures in those periods, respectively, among the set of common goods \mathbb{C}_t . p_{kt} and p_{kt-1} denote product k ’s average unit prices in those two periods; their log ratio is often called a log price relative.

The Tornqvist index has multiple attractive properties. First, as a “superlative” price index, it is closely related to other superlative price indices such as the Fisher and Sato-Vartia.³

³Every superlative price index is the change in the unit expenditure function (i.e., the exact price index)

We show that the Tornqvist and Sato-Vartia indices track each other closely in our empirical results. Second, as a superlative price index, the Tornqvist is also approximately consistent in aggregation, meaning that it is not sensitive to changes in product categorization or nesting strategies. Appendix 3.7.1 describes the interpretation of traditional price indices in more detail.

An important limitation of traditional price indices in the context of transactions data is that they are “matched-model” indices: they calculate price changes across the goods that were sold both in the base and in the current periods. Traditional indices therefore do not account directly for goods that enter or exit across periods, which we argue is an important source of changing product quality and is ubiquitous in item-level data. Another important limitation of traditional price indices in current practice is that statistical agencies’ data on sales and expenditure shares is often limited to disparate sources at higher levels of aggregation and lower frequency. For instance, the BLS uses expenditure shares from the Consumer Expenditure survey, with infrequently updated weights, to produce the Consumer Price Index (CPI). High-frequency scanner data connect the prices and quantities sold for each product, which allows for the construction of superlative price indices using internally consistent price and quantity data. We explore this advantage in our empirical analysis.

3.3.2 Hedonic Price Indices

Hedonic imputation allows a price index to account for product turnover by using product characteristics and an estimated hedonic relationship between characteristics and prices to impute the “missing” prices for entering and exiting products.

The log-level hedonic price model common in the literature takes the form

$$\ln p_{kt} = h_t(Z_k) + \eta_{kt}, \quad (3.2)$$

that is a second-order approximation for a wide class of utility functions in the absence of product turnover and taste shocks (Diewert, 1978). We generally discuss the Sato-Vartia in the context of the demand-based CES indices because it is exact for CES preferences under certain assumptions and because of our interest in contrasting it with other demand-based CES price indices.

where Z_k is a vector of observable characteristics for good k . The function $h_t()$ is often linear in parameters, and the hedonic equation is estimated with ordinary or weighted least squares regression. An important feature of equation (3.2) is that the hedonic function varies over time, i.e., the function $h_t()$ is estimated separately period-by-period. Underlying the hedonic approach is the assumption that utility can be specified as a function of the goods' characteristics. The time-varying estimation allows the hedonic function to capture changing consumer valuations, markups, or other changing aspects of market structure (Pakes, 2003).⁴

A core limitation of the log-level hedonic estimation approach outlined in equation (3.2) is that there are likely to be product characteristics that are relevant to the formation of prices but that the econometrician cannot observe. Erickson and Pakes (2011) introduce hedonic methods that can account for such unobserved characteristics. An important first step is to estimate hedonic models of price changes rather than price levels, e.g.

$$\Delta \ln p_{kt} = Z_k' \beta_t + v_{kt}. \quad (3.3)$$

This log-difference hedonic model estimates the change in hedonic price coefficients directly, which “differences out” any unobservable item-level characteristics that have a fixed influence on prices over time. This basic log-difference hedonic model does not, however, account for the influence of time-varying unobservable characteristics.

Erickson and Pakes (2011) therefore propose a modified approach that can account for time-varying unobservable characteristics. We call this approach the “TV approach” for short. Implementing the TV approach requires two steps. First, we estimate the log-level hedonic specification in equation (3.2) for period $t-1$. Second, we estimate a log-difference hedonic model including the lagged residuals from the first stage.⁵ The second estimating

⁴Although Pakes (2003) emphasizes that the estimated coefficients are not readily interpretable as marginal valuations of characteristics, the indices that emerge can be used as quality-adjusted estimates of changes in the cost of living.

⁵It can be shown that this characterization is equivalent to the time varying unobservables specification in Erickson and Pakes (2011). In that paper, they describe a closely related multi-step procedure. First, estimate the log levels hedonics and recover the residual. Second, estimate the log price relative on characteristics. Third, estimate the change in the residuals from the the log levels on the characteristics. Using the sum of the predictions from the latter two steps, as described in Erickson and Pakes (2011), is equivalent to using the predictions from equation (3.4).

equation is then

$$\Delta \ln p_{kt} = Z_k' \beta_t + \kappa_t \hat{\eta}_{kt-1} + v_{kt}. \quad (3.4)$$

Including the initial residuals from equation (3.2) in equation (3.4) allows the model to capture the influence of time-varying valuations of unobservable product characteristics to the extent that the initial residuals are correlated with price changes. In our analysis, we consider log-level, log first-difference, and TV approaches.

We also consider the related, but distinct, *time dummy* method that has been actively used in the research literature and by the BLS. We follow the recent literature (e.g., [Byrne et al., 2019](#)) using adjacent-period, weighted least squares estimation with Tornqvist market-share weights. Specifically, we pool observations from the adjacent periods $t-1$ and t and estimate hedonic regressions of the form

$$\ln p_{k\tau} = \alpha_{t-1,t} + \delta_t + Z_k' \gamma_{t-1,t} + \varepsilon_{k\tau}, \quad \tau = \{t-1, t\}, \quad (3.5)$$

where $\alpha_{t-1,t}$ is the constant, Z_k is the vector of characteristics for good k , $\gamma_{t-1,t}$ is a vector of estimated hedonic coefficients held fixed across periods $t-1$ and t , and δ_t is a fixed effect for period t .^{6,7} Exponents of the resulting coefficients δ_t can be interpreted as the quality-adjusted change in the price level between periods $t-1$ and t . Some limitations of the time dummy method are that it does not account for unobservable product characteristics and that it imposes constant coefficients on characteristics in adjacent periods. Appendix 3.7.2 provides additional discussion.

Our implementations of the TV approach and the time dummy method with the NPD data use standard econometric methods to estimate the hedonic function $h_t(\cdot)$. This approach is feasible with the NPD data because of the enormous value-added the NPD group provides in terms of item-level attributes.

⁶Letting T denote the total number of periods in the data, we estimate $T - 1$ separate pooled two-period regressions.

⁷We specify the hedonic regression equation (3.5) using the same vector of characteristics Z_k in each pair of adjacent periods. Occasionally, new features are introduced to the data. In pairs of adjacent periods entirely prior to the introduction of a new characteristic, it will be omitted from the regression because of collinearity with the intercept term. In pairs of adjacent periods in which the new feature is absent during period $t - 1$ and present during period t , the feature will be included in the estimated regression. Symmetric arguments apply for characteristics that exit.

The Nielsen data provided by the Kilts Center for Marketing at the University of Chicago does not contain pre-coded product attribute data for most products aside from short textual product descriptions. As outlined in Appendix 3.8.3, those descriptions are generally not coded to be human-intelligible.

To address these challenges, we have implemented deep neural networks to predict product prices and price changes from the product descriptions in the Nielsen Kilts Center data. Our approach parallels the TV approach of [Erickson and Pakes \(2011\)](#), in that it first predicts price levels and then, to capture time-varying unobservable effects, uses the prediction error in a second-stage neural net predicting price changes. This approach in the appendix and in a companion paper ([Cafarella et al., 2023](#)) that focuses on the machine learning methodology. In related work, [Bajari et al. \(2021\)](#) use an advanced machine learning approach that includes encoding image data as inputs into price predictions.⁸

We use weighted estimation methods in all of our hedonic indices, as is common in the literature using scanner data. We follow the recent time dummy literature by using Tornqvist expenditure weights, so that the time dummy method yields a quality-adjusted Tornqvist price index. We apply quantity-share weights for the estimation of the hedonic pricing functions in the hedonic imputation approaches using both traditional econometrics and ML methods. [Bajari et al. \(2021\)](#) also uses quantity-share weights in their implementation of ML methods for hedonic price indices using item-level transactions data. [Broda and Weinstein \(2010\)](#) and [Redding and Weinstein \(2020\)](#) advocate for quantity weighting based on the argument that unit values calculated based on a large number of purchases are better measured than those based on a small number of purchases. In Appendix 3.7.2, we also consider traditional regression-based TV hedonics using expenditure weights. Our main conclusions are broadly similar using both sets of estimation weights (see Table 3.A.3).⁹

In our main analysis using the TV approach, we assume the lagged residual for an entering good in the period prior to entry is zero. As a robustness check, in Table 3.A.1, we find very

⁸[Bajari et al. \(2021\)](#) provide novel methodology for encoding images via machine learning. They estimate hedonic models of price levels period by period.

⁹Note that we use the expenditure-share weights in the construction of the price indices themselves; the quantity-share weights are used only in estimation of the hedonic relationships in equations (3.2)–(3.4) and their machine-learning analogues.

similar results if we replace the predicted price relatives for entering goods with those from a hedonic regression that uses current period rather than lagged residuals and is otherwise equivalent to equation (3.4).¹⁰

We focus on full-imputation versions of the hedonic imputation indices that use predicted price relatives for all observations, including for common goods. Pakes (2003) shows that the hedonic Laspeyres imputation index bounds the exact change in the cost of living under a relatively weak set of assumptions. The key assumption is that consumers have preferences over the characteristics embodied in goods, rather than over the goods themselves. Indeed, full-imputation indices can be interpreted as characteristic price indices (Hill and Melser, 2006; De Haan, 2008). Using full-imputation indices also facilitates comparison with the time dummy method, as highlighted by De Haan (2008) and Diewert et al. (2008).¹¹ In addition, Erickson and Pakes (2011) argue that single- and double-imputation indices are subject to a form of selection bias, because they treat the hedonic estimation error for continuing, entering, and exiting goods in an asymmetric manner. Benkard and Bajari (2005), Diewert et al. (2008), and Bajari et al. (2021) have also used full-imputation indices. Our implementation of hedonic indices builds on and integrates the insights of this literature.

Hedonic price indices use the mapping between prices or price relatives and product characteristics among continuing goods to impute the “missing” prices or price relatives for entering and exiting goods. Note that characteristics turnover is distinct from product turnover. Indeed, we find that characteristics entry and exit rates are much smaller than the product entry and exit rates reported in Table 3.1. For boys’ jeans, occupational footwear and

¹⁰Erickson and Pakes (2011) do not face this issue because they consider only hedonic Laspeyres indices, which account for exiting goods but not entering goods. We also consider the difference between the traditional and hedonic Laspeyres using the TV method below. We find the differences are similar to the analogous differences using the Tornqvist indices. The Laspeyres indices have other limitations, but they are not sensitive to this assumption about the residual prior to product entry.

¹¹De Haan (2008) argues that, in the absence of unobserved characteristics, these indices are “strikingly similar.” Diewert et al. (2008) note the similarities and also derive the conditions under which they are identical. They note the full imputation approach is more flexible and in practice yields different results than the time dummy method. Neither of these papers highlights the importance of unobserved characteristics, as do Erickson and Pakes (2011). Incorporating the TV approach developed by Erickson and Pakes (2011) to address unobserved characteristics in the full-imputation indices produces additional advantages over the time dummy method. For these reasons, we favor the full-imputation TV approach of Erickson and Pakes (2011) in our hedonic indices.

memory cards we observed essentially zero characteristics entry over our sample period, although we do observe brand entry and exit (at rates less than 1 percent) for occupational footwear and boys' jeans. For headphones and coffee makers, we observe very low characteristics entry and exit rates (less than 0.1 percent), with slightly higher sales-weighted rates (as high as 0.5 percent for coffee makers). This evidence is consistent with the view that new characteristics diffuse slowly through the entry of new goods, and characteristics disappear from the available bundle slowly through product exit. Relatedly, new goods often have *more* of an important characteristic (e.g., size and speed of memory cards), while exiting goods often have *less* of those characteristics, so that product turnover involves upgrading of existing characteristics rather than the entry and exit of characteristics themselves.

3.3.3 Demand-Based Price Indices

In this section, we describe our use of exact cost-of-living indices for Constant Elasticity of Substitution (CES) demand systems. They provide tractable, implementable price indices that can account for quality change and product turnover. [Redding and Weinstein \(2020\)](#) characterize the unit expenditure function for a representative consumer with CES preferences can be characterized as

$$P_t = \left[\sum_{k \in \Omega_t} \left(\frac{p_{kt}}{\varphi_{kt}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}, \quad (3.6)$$

where $\sigma > 1$ is the consumer's elasticity of substitution between products, φ_{kt} is an appeal parameter for product k , and Ω_t is the set of products sold in period t . Both the set of products sold Ω_t and product-level appeal φ_{kt} may vary over time. Equation (3.6) is not empirically implementable, because the appeal parameters φ_{kt} are unobservable. The standard Sato-Vartia and Feenstra indices ([Sato, 1976](#); [Vartia, 1976](#); [Feenstra, 1994](#)) are also based on equation (3.6), but with the φ_{kt} restricted to have constant values over time (i.e., $\varphi_{kt} = \varphi_k$).

The Sato-Vartia index is exact for CES preferences in the absence of product turnover or time-varying product appeal. We denote product k 's expenditure share among all goods

sold in period t as s_{kt} and its expenditure share among common goods \mathbb{C}_t as s_{kt}^* . Letting c_{kt} be the quantity of good k purchased in period t , s_{kt} and s_{kt}^* are defined as

$$s_{kt} \equiv \frac{p_{kt}c_{kt}}{\sum_{l \in \Omega_t} p_{lt}c_{lt}} \quad \text{and} \quad s_{kt}^* \equiv \frac{p_{kt}c_{kt}}{\sum_{l \in \mathbb{C}_t} p_{lt}c_{lt}}. \quad (3.7)$$

The log Sato-Vartia index is then defined as

$$\ln \Phi_{t-1,t}^{SV} \equiv \sum_{k \in \mathbb{C}_t} \omega_{kt} \ln \left(\frac{p_{kt}}{p_{kt-1}} \right), \quad \omega_{kt} \equiv \frac{s_{kt}^* - s_{kt-1}^*}{\ln(s_{kt}^*) - \ln(s_{kt-1}^*)} \bigg/ \left(\sum_{k \in \mathbb{C}_t} \frac{s_{kt}^* - s_{kt-1}^*}{\ln(s_{kt}^*) - \ln(s_{kt-1}^*)} \right). \quad (3.8)$$

The [Feenstra \(1994\)](#) index generalizes the Sato-Vartia index to account for turnover in the set of goods sold Ω_t . We define the terms $\lambda_{t,t-1}$ and $\lambda_{t-1,t}$ as

$$\lambda_{t,t-1} \equiv \frac{\sum_{k \in \mathbb{C}_t} p_{kt}c_{kt}}{\sum_{k \in \Omega_t} p_{kt}c_{kt}}, \quad \lambda_{t-1,t} \equiv \frac{\sum_{k \in \mathbb{C}_t} p_{kt-1}c_{kt-1}}{\sum_{k \in \Omega_{t-1}} p_{kt-1}c_{kt-1}}. \quad (3.9)$$

The log Feenstra index is then defined as

$$\ln \Phi_{t-1,t}^{Feenstra} \equiv \frac{1}{\sigma - 1} \ln \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}} \right) + \ln \Phi_{t-1,t}^{SV}. \quad (3.10)$$

Letting $ER_{t-1,t}$ and $XR_{t-1,t}$ represent the sales-weighted product entry and exit rates as, the log Feenstra adjustment term can be approximated as $\ln \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}} \right)^{\frac{1}{\sigma-1}} \approx \frac{1}{\sigma-1} (XR_{t-1,t} - ER_{t-1,t})$. The Feenstra adjustment factor for product turnover (or ‘‘Lambda Ratio’’) thus indicates a downward adjustment to the Sato-Vartia index when the sales share of entering products is higher than the sales share of exiting products; it collapses to one in the absence of product turnover.¹²

The CUPI generalizes the Feenstra index to allow for time-varying product-level appeal. [Redding and Weinstein \(2020\)](#) emphasize that including time-varying product appeal is essential for the CES demand system to be consistent with the observed micro variation in prices and quantities, because quantities purchased often change even when relative prices do not. They specify a normalization on the changes in the appeal shocks so that there

¹²We use the actual Feenstra adjustment and not the approximation in our implementation.

is no change in geometric average tastes at the product group level for common goods. This assumption, combined with their assumption, which we also maintain, that consumers have Cobb-Douglas preferences across product groups, guarantees that product-level appeal shocks do not spill across product groups.

Redding and Weinstein (2020) derive an empirically implementable exact price index in this setting (the CUPI) as

$$\Psi_{t-1,t}^{CUPI} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}} \right)^{\frac{1}{\sigma-1}} \frac{\tilde{P}_t^*}{\tilde{P}_{t-1}^*} \left(\frac{\tilde{S}_t^*}{\tilde{S}_{t-1}^*} \right)^{\frac{1}{\sigma-1}}. \quad (3.11)$$

where we have followed Redding and Weinstein (2020) in using capital letters to indicate aggregates across product varieties, in denoting the geometric mean of a variable x as \tilde{X} , and in denoting the geometric mean over the set of common goods with an asterisk. \tilde{P}_t^* therefore represents the geometric mean of prices and \tilde{S}_t^* represents the geometric mean expenditure shares on common goods in period t . We call the CUPI's second term the “ P^* ratio” and its third term the “ S^* ratio.”

The log version of the CUPI is given by

$$\ln \Phi_{t-1,t}^{CUPI} = \frac{1}{\sigma-1} \ln \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}} \right) + \frac{1}{N_{C_t}} \sum_{k \in C_t} \ln \left(\frac{p_{kt}^*}{p_{kt-1}^*} \right) + \frac{1}{\sigma-1} \frac{1}{N_{C_t}} \sum_{k \in C_t} \ln \left(\frac{s_{kt}^*}{s_{kt-1}^*} \right), \quad (3.12)$$

where N_{C_t} denotes the number of common goods (i.e., products sold both in period t and in period $t-1$). Equation (3.12) clarifies that two of the CUPI's three terms are unweighted geometric means. This property is important for the CUPI's empirical implementation. Equation (3.12) also shows that the P^* ratio is simply the traditional Jevons index.

Each of these price indexes exactly recovers the change in the consumer's cost of living under different assumptions. The Sato-Vartia price index is exact if there is no product turnover and no time variation in product appeal.¹³ The Feenstra-adjusted Sato-Vartia index is exact in the presence of product turnover but the absence of time-varying product appeal. The CUPI is exact under the more general conditions of product turnover and time

¹³Feenstra and Reinsdorf (2007) show the Sato-Vartia index is unbiased in expectation with randomness in tastes under restricted conditions. Appendix 3.9.2 contains a further related discussion of this topic.

variation in product appeal. We find that these generalizations of the Sato-Vartia index are empirically relevant.

Although the CUPI is the most general of CES exact price indices that we consider, its inclusion of two unweighted geometric mean terms contrasts with the Sato-Vartia and Feenstra indices, which include only expenditure-weighted terms. Because the CUPI's unweighted terms are sensitive to products with very small expenditure shares, the CUPI can feature large measured price changes from what appear to be economically minor products.¹⁴

[Redding and Weinstein \(2020\)](#) adjust their empirical implementation of the CUPI by applying what we call a “common goods rule,” which defines the set of goods over which the P^* and S^* ratio terms are calculated (i.e., the goods included in the set \mathbb{C}_t). The goods excluded from the set of common goods are reallocated to the product turnover component (Feenstra adjustment factor), which is expenditure weighted. A common goods rule of this sort can be motivated by the argument that it takes time for goods to enter and exit the market. Consistent with this argument, [Redding and Weinstein \(2020\)](#) restrict the set of common goods in their empirical CUPI to those that are sold for a sufficiently long duration both prior to period $t-1$ and subsequent to period t .¹⁵ A limitation of this duration-based common goods rule is that it requires forward-looking information to implement, and thus it is not feasible to implement in real time. We find that we can mimic Redding and Weinstein's results using a purely backward-looking rule that can be implemented in real time. As we will see, the empirical effect of the S^* ratio varies significantly depending on the implementation of the common goods rule. We discuss potential reasons for the common goods rule's importance in Section 3.4.4 and Appendix 3.9 below.

¹⁴We thank Rob Feenstra for first bringing this point to our attention in his discussion of [Ehrlich et al. \(2022\)](#).

¹⁵[Redding and Weinstein \(2020\)](#) measure annual CUPI inflation from the fourth quarter of one year to the fourth quarter of the next year. Defining those quarters as periods $t-1$ and t , they define common goods as those sold in both of those quarters as well as in the 3 quarters prior to $t-1$ and the 3 quarters subsequent to t . In addition, they require the good be sold for at least 6 years total (although not necessarily consecutively).

3.4 Results

In this section, we present and discuss the traditional, hedonic, and demand-based exact price indices we have calculated in the item-level data. We focus first on our results from the NPD data, because the richness of the data permits more exploration of alternative methods.

3.4.1 NPD Results

Hedonics

We consider a wide variety of hedonic specifications as discussed in Section 3.3.2. We find that predicting log price changes directly produces significantly better model fit than estimating log price levels in periods $t-1$ and t separately and then forming a predicted log price change. Using the TV approach, which accounts for time-varying unobservable characteristics by including the lagged residual from a log-level regression, further increases the model fit across all product groups. The results therefore support the argument in [Erickson and Pakes \(2011\)](#) that estimating price changes helps to account for unobservable product characteristics, and that including the first-stage residual from predicting price levels in the estimation of log price changes provides a further advantage.

Table 3.A.4 provides diagnostics on the goodness of fit for our primary specifications. The column under the sub-header “Log Price Level” shows the average quarterly R-squared for products’ predicted log price levels. The average R-squared values range from 0.62 to 0.72 across product groups. The columns labeled “EP1” display results from predicting log price changes directly, without using the lagged residual. Predicting price changes is inherently a much more difficult task than predicting price levels, because the latter reflect cross-sectional differences in product characteristics, while the former reflect changes in the mapping between prices and characteristics over time. The average R-squared values using the “EP1” approach range from 0.09 to 0.47. Finally, the columns labeled “EP2” show results using the TV approach, which includes the lagged first-stage hedonic residual to predict log price changes. The average R-squared values using this approach range from 0.13 to 0.50 across product groups, an improvement on the results using the EP1

approach.¹⁶

Figure 3.2 compares several hedonic price indices for the five NPD product groups to the traditional Tornqvist index. We focus primarily on hedonic Tornqvist indices using fixed unobservables and time-varying unobservables approaches (the TV approach). For comparison we also present results from the time dummy hedonic approach. The values displayed in the figure are annual percent changes in the 4th quarter of each year from chained cumulative quarterly indices. The various price indices track each other broadly, but they also display some systematic differences. For all product groups, the TV approach yields a lower rate of price inflation compared to the traditional Tornqvist, the time dummy based index, or the first-difference based index. The gap between the traditional Tornqvist and the TV approach indices varies considerably across product groups, with the largest average differences for memory cards (-2.9 percentage points annually) and headphones (-2.5), and smaller differences for coffee makers (-0.70), boys' jeans (-1.30), and occupational footwear (-0.42).

The time dummy method suggests a notable hedonic adjustment for coffee makers relative to the traditional Tornqvist index, but for other products the difference is modest or is positive rather than negative. Our finding of limited quality adjustment the time dummy method is broadly consistent with the discussion in [Erickson and Pakes \(2011\)](#). As they emphasize, traditional hedonic approaches cannot account for the changing valuations of unobservable product characteristics, and in particular, how those changing valuations interact with product turnover. For example, if entering goods have desirable unobserved characteristics and correspondingly high prices, then the time dummy method may erroneously suggest a higher index value relative to the traditional Tornqvist.¹⁷ The previous literature has highlighted other limitations of the time dummy approach; for instance, [Pakes \(2003\)](#) raises

¹⁶The discussion above focuses on R-squared values using quantity-weighted estimation. Table 3.A.4 also shows results from expenditure-weighted estimation in the columns labeled “EW” for the weights. Quantity-weighted estimation yields higher R-squared for log price changes on average than expenditure-weighted estimation.

¹⁷For headphones, the traditional Tornqvist is notably lower in 2016 compared to the Hedonic Tornqvist using the time dummy method. This is a year when the share-weighted average price per item increases substantially. This pattern is consistent with entering goods having higher prices than existing goods. The time dummy method still yields a negative price change in that year, but not as negative as the standard Tornqvist. The hedonic Tornqvist TV method yields a more negative price index than the standard Tornqvist.

questions about the bound implied by the time dummy method, while [Diewert et al. \(2008\)](#) point out that the time dummy method requires more restrictive assumptions than the other hedonic approaches.

Figure 3.2 illustrates the contrasting results from using the fixed and time-varying unobservable estimation strategies.¹⁸ The hedonic Tornqvist indices using the fixed unobservables specification show inconsistent patterns of hedonic adjustment across product groups. In contrast, the hedonic Tornqvist using the TV method shows consistently greater deflation than the traditional index across product groups. This difference suggests that it is important to permit time-varying valuations of unobservable characteristics.

Figure 3.3 provides further evidence on the efficacy of the TV approach by displaying results with key observable characteristics left out of the hedonic estimation. Specifically, for memory cards the memory size is omitted, and for the other product groups, the large brand dummy variables are omitted. Omitting these informative characteristics from the estimation equation has a minimal effect on the resulting price indices. Appendix Figure 3.A.2 presents additional analyses showing that omitting those characteristics has a much larger effect on the hedonic indices using a log-level estimation approach.

Our results are consistent with the findings in [Erickson and Pakes \(2011\)](#). They present examples (e.g., for televisions) in which standard log-level hedonic estimation suggests higher rates of inflation than traditional matched models. They show, however, that using their methodologies to account for unobservable product characteristics (both using fixed valuation of unobservables and time-varying unobservables) produces systematically lower estimated inflation than the traditional matched models.

CES Demand-Based Price Indices

We turn now to CES demand-based price indices. The Feenstra index and the CUPI require estimates of the elasticities of substitution in their empirical implementation. Our baseline approach is to estimate a single elasticity for each of the NPD product groups. We employ

¹⁸To reiterate, the fixed unobservables strategy uses a log first-difference hedonic model, which the time-varying unobservables augments by including the residual from a first-stage hedonic estimate of log price levels.

the method used by [Feenstra \(1994\)](#) and [Redding and Weinstein \(2020\)](#) for this purpose.¹⁹ Table 3.2 reports the estimated elasticities, which range from about 5.2 to 7.8, consistent with the literature. The table also reports estimates from nested specifications, which we discuss below.

Figure 3.4 plots the Sato-Vartia, Feenstra, and CES unified (CUPI) price indices, as well as the components of the latter two indices. The baseline CUPI is calculated without a common goods rule and without any nesting within product groups. The Lambda Ratio and S^* ratio components in the figure are scaled by $\frac{1}{\sigma-1}$ so that the CUPI is the sum of the three components; see equation (3.11). We find that the CUPI shows low inflation relative to the Feenstra index and quite low inflation in absolute terms. In all goods but occupational footwear, the CUPI produces an estimate of 30%–40% annual declines in the price level annually, and it often falls 10–30 percentage points more quickly than the Feenstra index.

The large differences between the Feenstra index and the CUPI in these product groups arise from two sources. The first is the difference between the P^* ratio (Jevons index) and the Sato-Vartia index. The Sato-Vartia is a weighted average log price change among common goods, whereas the P^* ratio is an unweighted average. In boys' jeans, for instance, the P^* ratio is far below the Sato-Vartia. The difference between the weighted and unweighted log price ratios for common goods suggests there are a large number of low-share goods experiencing price declines that are driving down the CUPI. The second source is the introduction of the S^* ratio in the CUPI, intended to account for changing consumer tastes. Almost uniformly, the S^* ratio contributes a large downward shift to the CUPI. It is also an unweighted geometric mean that is sensitive to low-share goods.

The CUPI's sensitivity to low-share goods led [Redding and Weinstein \(2020\)](#) to introduce a common goods rule (or CGR) to the index. We implement a related but distinct methodology that can be implemented in real time using only current and backward looking information available in quarter t . For our NPD analysis, we specify a market share threshold for goods present in periods t and $t-1$ to be considered as common goods for the Jevons

¹⁹This method double-differences the demand and supply curves sweeping out time and product group effects. The double-differenced demand and supply shocks are assumed to be uncorrelated but heteroskedastic across products. This yields a GMM specification for estimation. As in [Redding and Weinstein \(2020\)](#), the weighting matrix is based on quantity-weights.

and the S^* ratio terms of the CUPI.²⁰ Goods below this threshold are excluded from the set of common goods, but they still enter the CUPI through their inclusion in the Feenstra adjustment term (lambda ratio). We consider alternative market share percentile thresholds in our analysis of the NPD data.²¹

Figure 3.5 illustrates the CUPI's sensitivity to the CGR for different market share thresholds. We consider market share thresholds for continuing goods in t and $t-1$ of the 10th, 30th, and 50th percentiles. We depict the CUPI for these different CGRs alongside the Feenstra index and the CUPI without a CGR. Implementing the restriction on the set of "common goods" by market share raises the CUPI by cutting off the low end of the market share distribution from unweighted relative comparisons and shifting it to the weighted entry/exit adjustment term. In that sense, applying a stricter CGR moves the CUPI closer to the Feenstra-adjusted Sato-Vartia index, which combines a traditional matched model index with an adjustment for entry and exit.

The resulting price indices generally shift up as successively stricter definitions of common goods are imposed. For some product groups, such as memory cards, the CUPI using the CGR at the 30th or 50th percentile yields inflation measurements similar to the Feenstra index. For products groups such as headphones and boys' jeans, however, the CUPI shows noticeably lower inflation than the Feenstra index even using with a 50th-percentile CGR threshold (i.e., excluding half of products from the set of common goods).

The key takeaway from this analysis is that the CUPI is sensitive to the specific definition of the CGR, and that sensitivity varies across product groups. In contrast to the finding in Redding and Weinstein (2020) that the CUPI eventually stabilizes as successively stricter duration-based CGRs are applied, we do not find evidence that the CUPI stabilizes as stricter share-based CGRs are applied. A 50th-percentile threshold for the market share of goods present in t and $t-1$ implies that an entering good does not count as a common good until it reaches the top half of the market share distribution. Similarly, a good that

²⁰The details of the procedure are as follows. Compute the X th percentile of the expenditure shares within product groups in both period $t-1$ and period t . A common good must exceed the X th percentile in both periods.

²¹In our analysis of the Nielsen data, which is a longer panel, we consider further alternative approaches to define common goods. In our analysis of chain drift below, we also consider the impact of the CGR implemented over a longer horizon in the NPD data.

is on its way to exit and that falls below the 50th percentile of market share is put into the entry/exit group (and becomes part of the Feenstra adjustment term).

We are sympathetic to the view that some form of CGR is a sensible and necessary component of empirically implementing the CUPI. The primary inference we draw from our own analysis and the literature to date is that the CUPI is sensitive to the specification of the CGR, and more research is necessary on best empirical practice in implementing the index. Further research into the dynamic process of the entry and exit of goods should be a part of such research. Our analysis in Figure 3.1 is a step in that direction. It is likely that process varies by product group, consistent with our results showing the CUPI's differential sensitivity to various CGRs across product groups.

[Martin \(2020\)](#) notes that the S^* ratio can reflect not only shifting preferences, but also any model misspecification, including a nested preference structure. The CUPI's assumed CES preference structure imposes an equal elasticity of substitution within product groups, and violations of this assumption could lead to biased measures of inflation. Furthermore, the CUPI is more vulnerable to this issue than the other CES price indices we consider.²²

We have explored nesting products into subgroups in the NPD data, but we have consistently found that nesting does not overturn the CUPI's extremely negative inflation readings relative to the Feenstra index. We describe our analysis in detail in Appendix 3.9.1. To summarize, we have explored two different approaches to nesting, a characteristics-based approach and an approach based on predicted prices using hedonic regressions. Neither nesting method materially changes the inflation measurements of the CUPI, suggesting that aggregation issues are not the primary driver of its estimates of rapid deflation in the product groups we consider.

²²More precisely, [Martin \(2020\)](#) shows that the CUPI is not consistent in aggregation. [Vartia \(1976\)](#) defines consistency in aggregation as the equality of a single-stage or two-stage index number. In the single-stage of an index number, all goods are included in a single aggregation. In a two-stage construction, the index is computed for a number of subgroups, and the subgroups are aggregated using the same index number formula. [Diewert \(1978\)](#) shows that the Sato-Vartia index is consistent in aggregation.

Comparing Traditional, Hedonic, and Exact Price Indices

Figure 3.6 presents the main traditional, hedonic, and demand-based price indices that we have considered for all five product groups. Because the CUPI indices are outliers for some groups, Figure 3.7 displays price indices without the CUPIs but with the addition of the Laspeyres index. Figures 3.8 and 3.A.3 in Appendix 3.9.4 present plots of chained price index levels calculated by chaining the quarterly price indices underlying Figures 3.6 and 3.7. They therefore illustrate the cumulative effects of the differences between the various indices. Likewise, Table 3.3 reports the chained index levels in 2018:4, reflecting the cumulative price changes since 2014:4, when all indices are normalized to one.

The price indices follow a roughly similar pattern of relative orders across these product groups. Figure 3.7 shows that the Laspeyres index typically shows the least deflation. The traditional Tornqvist and Sato-Vartia tend to track each other closely and to show more rapid deflation than the Laspeyres, as expected given that they are both superlative price indices that account for substitution. The Feenstra and hedonic Tornqvist using the TV method in turn tend to show greater deflation than their unadjusted counterparts, indicating the importance of product entry and exit. Finally, in Figure 3.6, the CUPI (both baseline and nested by product characteristics) shows the greatest deflation, especially for headphones and boys' jeans. The substantial gap between the CUPI and the Feenstra index in headphones and boys' jeans is especially striking given our imposition of a 30th-percentile CGR.

The gap between the traditional Laspeyres and Tornqvist indices for most product groups highlights the advantages of using item-level scanner data, which permits construction of a superlative price index with internally consistent prices and expenditure shares in adjacent periods. The gap is especially striking in the cumulative price indices shown in Figure 3.A.3. The cumulative gaps are on the order of 5–10 percentage points from 2014 to 2018 in coffee makers, occupational footwear, and boys' jeans, and significantly larger in the “high-tech” categories of memory cards and headphones. The gap also varies over time, consistent with the Laspeyres index exhibiting a time-varying substitution bias. Thus, using scanner data can produce substantial improvements in price measurement even without performing quality adjustment.

Figure 3.A.3 also shows that the hedonic Tornqvist using the TV method tends to indicate larger cumulative quality adjustment via product turnover than the Feenstra index. The Feenstra index indicates approximately 2 percentage points of cumulative disinflation beyond the Sato-Vartia index across all five product groups. The hedonic Tornqvist indicates roughly the same adjustment (relative to the traditional Tornqvist) for coffeemakers and occupational footwear, but larger adjustments, of roughly 5–7 percentage points, in memory cards, headphones, and boys’ jeans.

3.4.2 Nielsen Results

In the main text, we restrict our analysis of results for the Nielsen scanner data to the food product groups. For the CES exact price indices, our empirical implementation in the Nielsen data largely follows our strategy in the NPD data. The Feenstra index and the CUPI require estimates of elasticities of substitution within product groups. As in the NPD data, we use the [Feenstra \(1994\)](#) procedure to estimate those elasticities. The estimated elasticities for the 50+ product groups in food display considerable variation. The median elasticity is about 6, the 10th percentile is about 4, and the 90th percentile is 12. These patterns are similar to those reported in [Redding and Weinstein \(2020\)](#).

For the hedonic TV approach, we combine the insights of [Erickson and Pakes \(2011\)](#) with the machine learning approach developed in our companion paper, [Cafarella et al. \(2023\)](#), and summarized in Appendix 3.8.5. The machine learning approach allows us to exploit the Nielsen scanner data’s unstructured information on item-level attributes.

The median in-sample R-squared of the hedonic price predictions is roughly 85% for the food product groups and nearly 90% for the nonfood groups.²³ The median out-of-sample R-squared is roughly 75% for food and 75% for nonfood. The model’s predictive performance is comparable to that of [Bajari et al. \(2021\)](#), who report out-of-sample R-squared values of 80–90% in their best-performing specifications using the rich product text and image information in their data set. For log price changes, our median in-sample R-squared is above 50% for the food product groups and above 40% for the nonfood product groups. The median out-of-sample R-squared values decline to nearly 20% and below 10% for

²³See [Cafarella et al. \(2023\)](#) for details.

the food and nonfood product groups, respectively. We consider our procedure to be very successful in light of the limited attribute information available in the data set.

We again explore alternative CGRs to calculate the CUPI. The Nielsen data provides a longer panel than the NPD data, which allows the exploration of alternative CGRs that depend on the duration of goods' time in the market to date. We implement a modified approach to defining the CGR rule in the Nielsen data as follows. We first compute percentiles of the pooled sales distribution within a narrow product group for pooled sales in periods $t-1$ and t . Common goods are defined as goods sold in both periods, and which have sales in period t above the X th percentile of this pooled sales distribution. This alternative approach to defining the CGR allows us to consider longer duration-based alternatives.²⁴

Figure 3.9 shows the results for the aggregated food categories of the CUPI and its components using various CGRs defined by different sales-based percentiles. The Feenstra adjustment and S^* ratio terms have again been scaled by $\frac{1}{\bar{\sigma}-1}$ for each constituent product group so that the components sum to the CUPI. The Feenstra adjustment (lambda ratio) and Jevons index (P^* ratio) components of the CUPI show very little sensitivity to the alternative CGRs. Indeed, the plots for the different values are nearly indistinguishable. In contrast, the S^* ratio is very sensitive to the CGR in the Nielsen data, which leads directly to sensitivity in the CUPI.²⁵ The baseline CUPI without a CGR percentile threshold has average four-quarter price inflation about 10 percentage points below the Feenstra. Using a 50th percentile for the CGR yields a price index that is much closer to the Feenstra index.

We consider alternative specifications of the CGR using market thresholds using percentiles of sales pooled over over the current and prior 4 quarters. In addition and critically, a common good is defined in this context if it is present in periods t and $t - 4$. Using a duration component in the CGR puts more weight on goods present for the longer horizon, yielding greater comparability with the duration-based CGR used by Redding and Weinstein (2020).²⁶ Appendix 3.8.4 shows that using this longer horizon approach for computing

²⁴In unreported results, we have found that the Nielsen results using the identical CGR used in the NPD data yields very similar results to those reported here using a two-quarter horizon.

²⁵Recall that the S^* ratio is an unweighted geometric mean, which is sensitive to small market shares.

²⁶An advantage of this alternative duration-based CGR for the purposes of producing real-time statistics is that it does not require forward-looking information.

sales percentiles, a CGR with a 10th-percentile sales threshold yields results comparable to a CGR with a sales threshold between the 25th and 50th percentiles using a two quarter horizon.²⁷

Figure 3.10 presents a full set of price indices for the Nielsen scanner food product groups in change and level forms.²⁸ The panels of the figure include the BLS CPI computed for the same Nielsen product groups.²⁹ We find that the CPI and the traditional Laspeyres index track each other closely in Nielsen’s food product groups for the first part of the sample period, with a discrepancy arising towards the end of the period. The Tornqvist and Sato-Vartia indices are lower and track each other closely. The quality-adjusted indices (Feenstra, hedonic Tornqvist using the TV approach, and CUPI) are even lower.

The cumulative level implications highlight that the hedonic Tornqvist is about 4 percentage points lower in 2015 than the traditional Tornqvist, and the Feenstra index is about 5 percentage points lower than the Sato-Vartia. These substantial cumulative differences for the food product groups suggest that quality improvement via product turnover has not been limited to products where technological progress is most visible. It is also striking that these two distinct relative comparisons yield such similar quantitative implications.

Using a 25th-percentile CGR, the CUPI is more than 40 percentage points lower than the Feenstra index in 2015; using a 50th percentile CGR reduces the difference to 20 percentage points. Alternatively, using the longer $t - 4$ to t horizon described above, the 10th-percentile CGR yields a difference of about 25 percentage points.³⁰

We consider the patterns in the Nielsen data to be broadly similar to the patterns in the NPD

²⁷Appendix 3.8.4 also explores the use of the Nielsen Consumer Panel and CGR sales-based percentile rules. Using the Consumer Panel enables us to more readily compare our results to those in Redding and Weinstein (2020).

²⁸For all indices, we aggregate across product groups using a Tornqvist aggregator with Divisia-style product group market share weights.

²⁹We thank the BLS for producing these calculations.

³⁰Results for the nonfood product groups, described in Appendix 3.8.2, show substantially greater departures between the BLS CPI and the Nielsen Laspeyres consistent with our concerns about the Nielsen scanner data’s representativeness for the nonfood product groups. The CUPI for nonfood is very low. With a 30th-percentile CGR, the CUPI price level (indexed to 2006) is almost 70 percentage points lower in 2015 compared to the Feenstra (the difference shrinks slightly to 40 percentage points with a 50th-percentile CGR). These results may arise partly from the limited coverage of nonfood items in the Nielsen scanner data.

data. Quality adjustment, either via hedonic approaches or the Feenstra product turnover adjustment, imparts a substantial downward adjustment on price indices. The CUPI suggests an even larger quality adjustment, but we note again its sensitivity to the CGR. This sensitivity manifests across alternative approaches to defining the CGR thresholds for common goods.

3.4.3 Chain Drift

A potential challenge to using transactions data to compute price indices is chain drift. This issue is particularly problematic with high-frequency indices computed from local transactions data (e.g., [De Haan and Van Der Grient, 2011](#)). Our analysis uses national data at a quarterly frequency, which mitigates this issue. Given our focus on comparing alternative approaches for computing price indices, we consider whether GEKS-type indices ([Gini, 1931](#); [Eltetö and Köves, 1964](#); [Sculz, 1964](#)) preserve the implications of our core findings.

Primarily, we follow [Bajari et al. \(2021\)](#) by computing a GEKS-type index (which we denote “GEKS-lite”) that is the geometric mean of the chained year-over-year index for the 4th quarter of each year and the directly computed (unchained) year-over-year price index for the 4th quarter. GEKS-lite offers an easily implementable alternative to a full GEKS procedure, which involves computing price indices over many possible horizons. Given that we are including hedonic indices which require re-estimation of models for alternative time horizons, the computational burden of implementing a full GEKS procedure is substantial. [Table 3.4](#) reports average annual chained and GEKS-lite indices for the five NPD product groups and alternative indices. The GEKS-lite price change index tends to show slower deflation than the chained price index for traditional price indices, but the differences are not large. For the hedonic indices, the GEKS-lite indices actually show faster deflation in three out of the five product groups. Among the demand-based indices, the GEKS-lite indices typically show slower deflation, but again these differences are modest quantitatively. The GEKS-lite CUPI shows substantially less deflation, but this difference also reflects the effects of applying a common good rule over a longer horizon.³¹

³¹The longer horizon affects the CGR because for the year-over-year measure the good must not only be above the Xth percentile in the appropriate samples, but also be present in quarters t and $t - 4$, as opposed to quarters t and $t - 1$.

The key result of this analysis is that applying the GEKS-lite procedure does not change the rank ordering of the various indices we have considered. The Laspeyres index yields higher inflation than the Tornqvist, which in turn is higher than the hedonic Tornqvist. Likewise, the Sato-Vartia yields higher inflation than the Feenstra, which in turn is higher than the CUPI. Notably, the traditional Tornqvist and Sato-Vartia (both superlative indices) are more sensitive to chain drift than key quality-adjusted alternatives (hedonic Tornqvist and Feenstra).

We also implement the rolling-year GEKS method (Ivancic et al., 2011) as a point of comparison. Given that this procedure is computationally burdensome, we only implement this for traditional price indices. In appendix Table 3.A.2, we show that results for rolling year GEKS are similar to those using the GEKS-lite approach.

Table 3.5 reports analogous chained and GEKS-lite indices for the aggregated food indices, which we generate following the same procedure as in prior sections.³² The table reports average annual indices for both specifications. The results for Nielsen’s food product groups show that we obtain similar, albeit slightly higher, rates of average inflation using the GEKS-lite compared to the chained indices. This pattern is especially noticeable for the CUPI, a result which again reflects the effects of applying the CGR over a longer horizon. Importantly, the rank ordering and the quantitative differences across alternative indices are preserved using the GEKS-lite based indices. Focusing on the GEKS-lite indices, inflation in Nielsen’s Food product groups is higher using the Laspeyres index than the Tornqvist, higher using the Sato-Vartia than the Feenstra, and higher using the Feenstra than the CUPI.³³

3.4.4 Taking Stock

The current system of price measurement implemented by the BLS samples a relatively small quantity of goods and then aims to follow them for a number of periods. When

³²That is, we compute the indices at the product group level and then use Divisia weights to aggregate to the food level.

³³We do not report the hedonic indices using the GEKS-lite procedure for Nielsen’s food product groups because of the large computational burden that would be required to apply our machine learning procedure to additional comparison periods.

goods disappear, the BLS uses various approaches to account for turnover, which often involve expert judgment. The BLS uses hedonic methods on a case-by-case basis for a small fraction of the market basket. Thus, while procedures to address quality changes and product turnover do currently exist, they are designed around a data collection and processing architecture based on hand-collection of price quotations and non-uniform approaches to addressing quality change. These procedures do not readily scale and do not take advantage of data housed in the information systems of businesses.

Incorporating transactions data into official statistics requires making several methodological decisions, such as whether to use traditional, hedonic, or exact price indices, and how to implement the chosen index. Putting aside the important issues of quality change and product turnover, traditional matched model price indices constructed from item-level data possess several advantages relative to the current system: the expenditure shares from the item-level data are internally consistent with the price data, and they are also available in real time. The data therefore permit the construction of superlative price indices such as the Tornqvist in real time. We find that the Tornqvist index tends to measure systematically lower inflation than the Laspeyres, with the gap varying over time and product groups.

If the item-level data contain information on product attributes, as they commonly do, hedonic methods can also be applied at scale in real time. We have found that the most robust approach for implementing hedonics at scale is to use the time-varying unobservables approach from [Erickson and Pakes \(2011\)](#). Our results provide ample support for their argument that it is important to correct for the reevaluation of the unmeasured characteristics of continuing, entering, and exiting goods. Our approach also extends their results by demonstrating that their methodology can be implemented at scale using transactions data and superlative price indices with internally consistent prices and expenditure weights.

Demand-based indices offer a useful alternative for comparison to hedonic indices. These indices are exact under certain sets of assumptions, and in the most general case (the CUPI), they can account both for quality change via product turnover and for time-varying product appeal for continuing goods. [Redding and Weinstein \(2020\)](#) argue that neglecting the latter issues can bias cost-of-living price indices.

The limitation of the CES demand-based approaches we have considered is their sensitivity to the strong assumptions of their underlying models, which may omit empirically important market imperfections. A central assumption of these approaches is the existence of a unified national market where all goods are available. In Appendix 3.8.6, we assess the realism of this assumption in the Nielsen data. We show that most products have less than 20 percent penetration across Nielsen metro areas. Items with greater sales volumes are sold across more areas, so the national market assumption is more realistic on a sales-weighted basis. Nonetheless, even on a sales-weighted basis, a distinct minority of items reach a truly national market.

We believe the failure of the national market assumption is likely to have a much larger effect on the CUPI than on the other price indices we have considered. Superlative price indices such as the Tornqvist and Sato-Vartia are approximately consistent in aggregation (Diewert, 1978), so the failure of the national market assumption is less troubling for those indices; a similar argument applies to the hedonic Tornqvist index. The Feenstra index generalizes the Sato-Vartia index with an expenditure-weighted term to correct for product turnover, so it also contains only expenditure-weighted terms. In contrast, the CUPI contains multiple unweighted terms, which mean that goods with small expenditure shares can have an outsized effect on the index.

In Appendix 3.9, we examine the behavior of the CUPI empirically, analytically, and in simulations under various assumptions about preferences and market structures. A few key results emerge from our examination. First, as noted above, in Appendix 3.9.1 we examine the CUPI's sensitivity to assumptions about nested preference structures. The various nesting structures we have explored, using characteristics-based and predicted price-based nests, do not meaningfully change the pattern that the CUPI measures significantly lower inflation than the Feenstra index even with the application of a stringent common goods rule.

Next, we present an analytical argument in Appendix 3.9.2 that the presence of time-varying product appeal does not, on its own, produce a bias in the Sato-Vartia and Feenstra indices. We show that, although the force that Redding and Weinstein (2020) argue will tend to impart an upward bias to those indices does exist, if appeal shocks are indepen-

dently and identically distributed (i.i.d.) over time, there is a symmetrical and offsetting force that will impart a downward bias. We present simulation evidence to support our analysis in Appendix 3.9.3. The simulations show that, although time-varying product appeal does not produce an average bias in the Sato-Vartia index in the presence of i.i.d. product appeal shocks, the Sato-Vartia is noisier than the CUPI in the presence of appeal shocks. The simulation results also show that rising dispersion in product appeal will cause the CUPI to measure lower inflation than the Sato-Vartia, which will be upward-biased. We show in Figure 3.A.5, however, that rising product dispersion appears empirically unable to account for the ubiquitously low inflation rates measured by the CUPI. Finally, we present simulation evidence from cases of market imperfections such as localized markets or product stockouts that may drive a downward bias in the CUPI, and that a CGR can help to ameliorate the bias in those cases. These simulations suggest that the failure of the national market assumption documented in Appendix 3.8.6 may be an important driver of the empirical behavior of the CUPI.

Taking stock of the empirical and theoretical evidence, we believe that the hedonic methods we have explored provide a sensible and data-driven approach to quality adjustment that can be applied at scale in transactions-level data. We believe that the demand-based indices that incorporate quality adjustment—particularly the Feenstra—provide useful benchmarks that should be used for purposes of comparison with the hedonic and traditional indices. On the other hand, current implementations of the CUPI involve strong assumptions about market structure and preferences that can yield anomalous results when taken to the data. Although modifications such as a CGR can ameliorate or eliminate those anomalies, there is limited theoretical or empirical guidance on what an appropriate correction might be. This topic should be a high priority for future research. As this paper demonstrates, item-level transactions data with prices, quantities and attributes permits exploration and comparison of a wide range of potential price indices, which should facilitate such future research.

3.5 Conclusion

Item-level transactions data with price, quantity, and attribute information enable considerable advances in the production of price indexes. These include price indexes that are

granular with respect to product, geography, and frequency; available in close to real time; adjusted for substitution across goods using superlative index formulas, and adjusted for quality change and rapid turnover of goods. The availability of contemporaneous price and quantity data and of rich item descriptions is fundamental to implementing these innovations. While the increased availability of “big data” therefore has considerable promise for improving price measurement, there are also considerable challenges to the production of quality-adjusted price indices. In particular, because of the rapid turnover of products in item-level data, quality change must be addressed at scale.

We address the challenge of implementing price indexes in big data by exploring and evaluating two alternative approaches for quality adjustment at scale: hedonic methods and demand-based methods. We find that it is important for hedonic methods to account for time-varying changes in the valuation of goods’ unobservable characteristics. We do so using the methodology of [Erickson and Pakes \(2011\)](#). We demonstrate that it is possible to implement hedonics at scale by using machine learning approaches. Our results show that traditional matched-model methods substantially overstate the rate of inflation. We find that these patterns are pervasive, that is, not limited to goods such as electronics where technological progress is most visible. Hence, the practice of statistical agencies of focusing hedonic adjustment in sectors where technological change is most visible can obscure pervasive quality change.

Among the demand-based methods we have considered, we find that the [Feenstra \(1994\)](#) index produces a relatively stable magnitude of quality adjustment across product groups that is also broadly consistent with the implications of our preferred hedonic methods. The CUPPI developed in the path-breaking work of [Redding and Weinstein \(2020\)](#) implies very rapid deflation without the imposition of a common goods rule. Our analysis suggests the need for further research on how to choose such a rule while preserving the CUPPI’s theoretical appeal and suggests some potential paths forward on that topic.

Finally, and importantly, item-level transaction data allow superlative indexes to be constructed in real time using internally consistent prices and quantities. We find that accurately measuring substitution effects among continuing goods with such data has a large effect on estimated inflation.

This paper is a step in demonstrating that using item-level transactions data at scale can lead to a re-engineering of key national indicators. It shows that accounting both for substitution effects and for quality change substantially lowers the estimates of inflation rates across a wide range of goods. These innovations therefore should have important implications for understanding the average rate of price change, with further implications for estimates of the rate of growth of output and productivity.

Table 3.1: Rates of Product Turnover: NPD Data

	Entry Rate		Exit Rate	
	All	Initial	All	Final
Memory Cards	5.8%	3.0%	6.0%	3.3%
Coffee Makers	5.7%	3.4%	4.5%	2.1%
Headphones	6.4%	3.8%	5.5%	2.9%
Boys' Jeans	11.5%	8.3%	7.8%	4.3%
Occupational Footwear	13.5%	9.1%	10.6%	5.5%

Notes: Average quarterly rates of product turnover. Entry/exit rates are computed as the number of entering/exiting goods as a percentage of common goods in the previous period. “Initial” entries are those for which the product was never observed in the data prior to the quarter. “All” entries include entries in which the product was previously observed prior to a spell of absence and the re-entered the data (i.e., “re-entries”). “Final” exits are those for which the product was never again observed in the data after the quarter. “All” exits include exits for which the product is subsequently observed after a temporary spell of absence (i.e., “temporary exits”). Transition quarter between data vintages excluded. Data come from NPD Group.

Table 3.2: Estimated Elasticities of Substitution: NPD Data

Product	Elasticity of Substitution
Headphones	7.634 (0.748)
Memory Cards	5.623 (0.484)
Coffeemakers	5.183 (1.289)
Occupational Footwear	7.31 (0.533)
Boys' Jeans	7.861 (0.565)

Notes: Elasticities of substitution for Feenstra index and CUPI estimated using approach of [Feenstra \(1994\)](#) and [Redding and Weinstein \(2020\)](#). Standard errors in parentheses. Data come from NPD Group.

Table 3.3: Alternative Price Indices, Levels in 2018q4 Relative to 2014q4: NPD Data

	Memory Cards	Coffeemakers	Headphones	Boys' Jeans	Occupational Footwear
Laspeyres	0.539	0.749	0.605	0.735	0.887
Hed. Laspeyres,TV	0.414	0.683	0.494	0.709	0.859
Tornqvist	0.467	0.688	0.607	0.726	0.872
Hed. Tornqvist,TV	0.399	0.666	0.541	0.680	0.856
Sato-Vartia	0.481	0.706	0.602	0.773	0.879
Feenstra	0.469	0.685	0.582	0.749	0.857
CUPI, CGR 30p	0.389	0.625	0.332	0.181	0.777
CUPI-N, CGR 30p	0.367	0.640	0.349	0.173	0.780

Notes: Values are cumulative changes in 2018:4 relative to the 2014 price level, with 2014 price level set to 1. CUPI-N is nested CUPI using characteristics approach. Data come from the NPD Group.

Table 3.4: Alternative Price Change Indices, Chained (C) vs GEKS-Lite (GL), NPD Data

	Memory Cards	Coffeemakers	Headphones	Boys' Jeans	Occupational Footwear
Laspeyres (C)	-13.89	-6.87	-11.74	-7.36	-2.92
Laspeyres (GL)	-12.16	-5.63	-11.77	-6.11	-2.15
Tornqvist(C)	-16.90	-8.86	-11.58	-7.63	-3.35
Tornqvist(GL)	-15.41	-6.64	-11.55	-5.56	-2.31
Hed.Tornqvist,TV(C)	-19.83	-9.56	-14.13	-8.93	-3.77
Hed.Tornqvist,TV(GL)	-20.6	-10.06	-14.51	-7.94	-3.76
Sato-Vartia(C)	-16.24	-8.24	-11.75	-6.20	-3.14
Sato-Vartia(GL)	-14.32	-6.36	-11.34	-4.13	-2.10
Feenstra(C)	-16.78	-8.92	-12.47	-6.92	-3.76
Feenstra(GL)	-16.46	-9.43	-13.06	-5.51	-3.80
CUPI,CGR 30p(C)	-20.64	-11.05	-24.08	-34.74	-6.08
CUPI,CGR 30p(GL)	-19.94	-9.41	-22.55	-26.91	-5.20

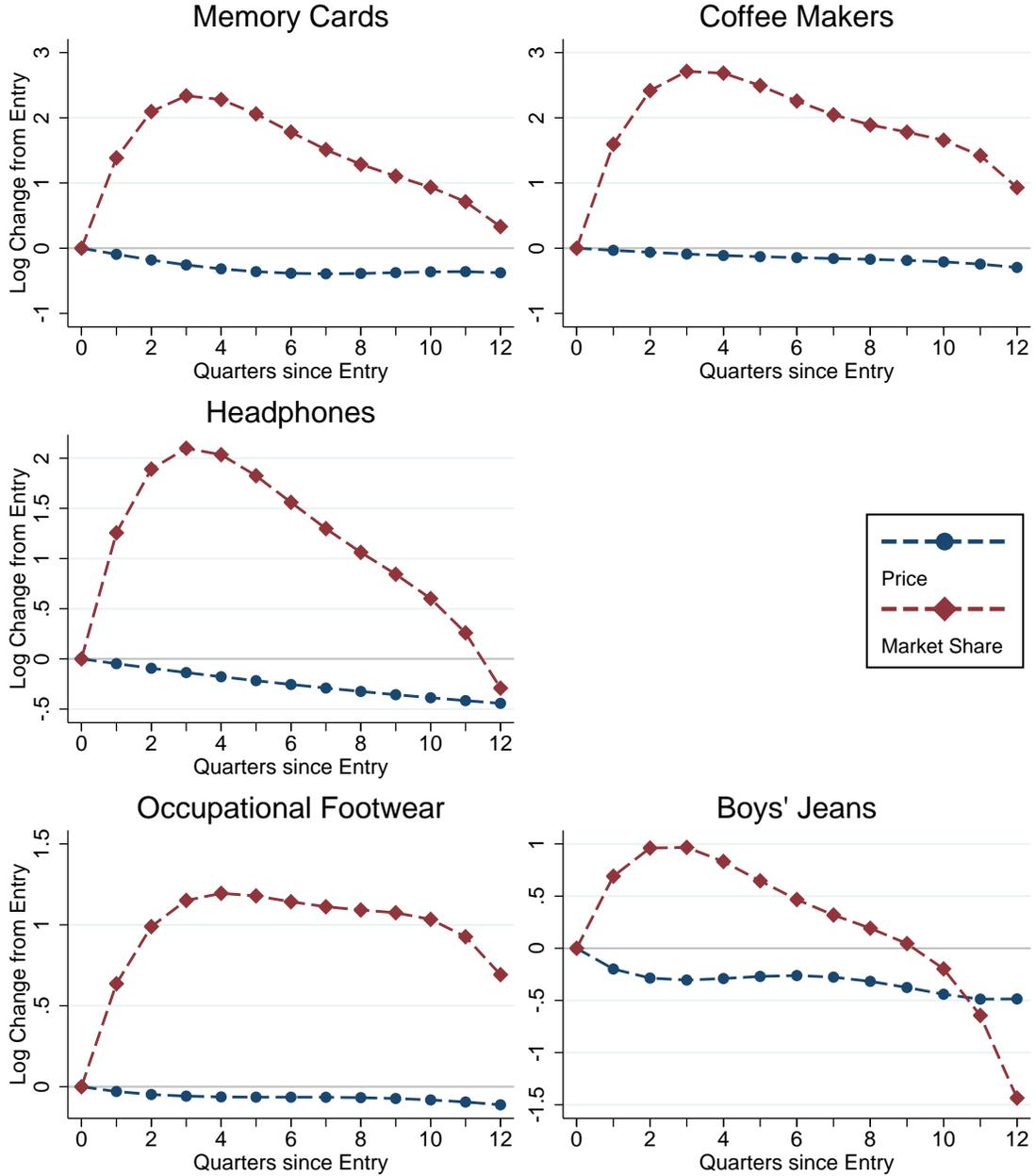
Notes: Chained values are averages of cumulative quarterly rates for year. GEKS-lite is the average of the geometric mean of the chained values and the YoY price indices for q4 for each year. Data come from the NPD Group.

Table 3.5: Alternative Price Change Indices, Chained vs GEKS-Lite, Nielsen, Food

Index	Chained	GEKS-Lite
Laspeyres	.014	.014
Tornqvist	.005	.009
Sato-Vartia	.007	.010
Feenstra	.003	.005
CUPI	-.034	-.020

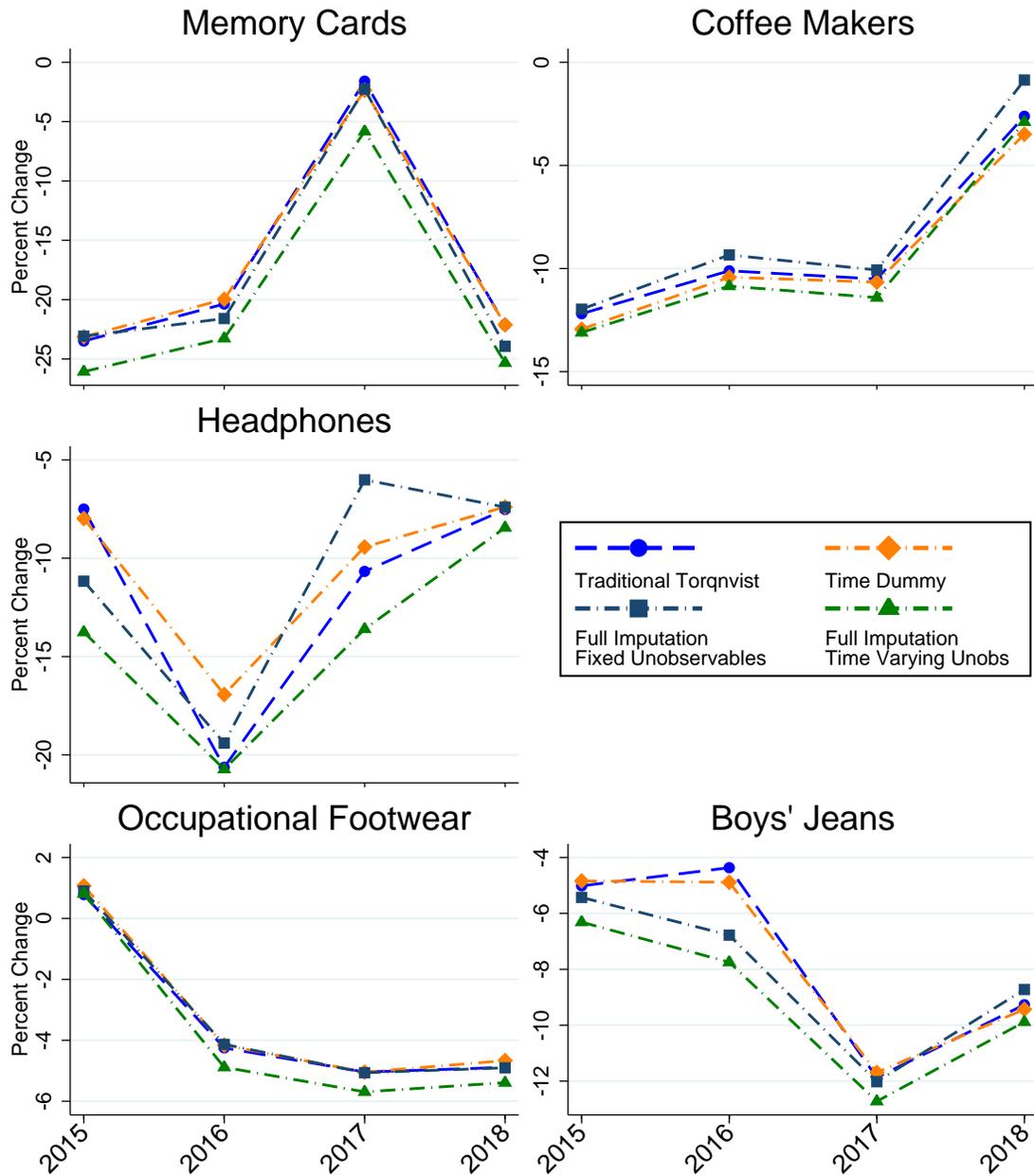
Notes: Chained values are averages of cumulative quarterly rates for year. GEKS-lite is the average of the geometric mean of the chained values and the YoY price indices for q4 for each year. Laspeyres is the geometric Laspeyres. CUPI uses 25th percentile CGR. Data come from Nielsen.

Figure 3.1: Product Lifecycle Dynamics



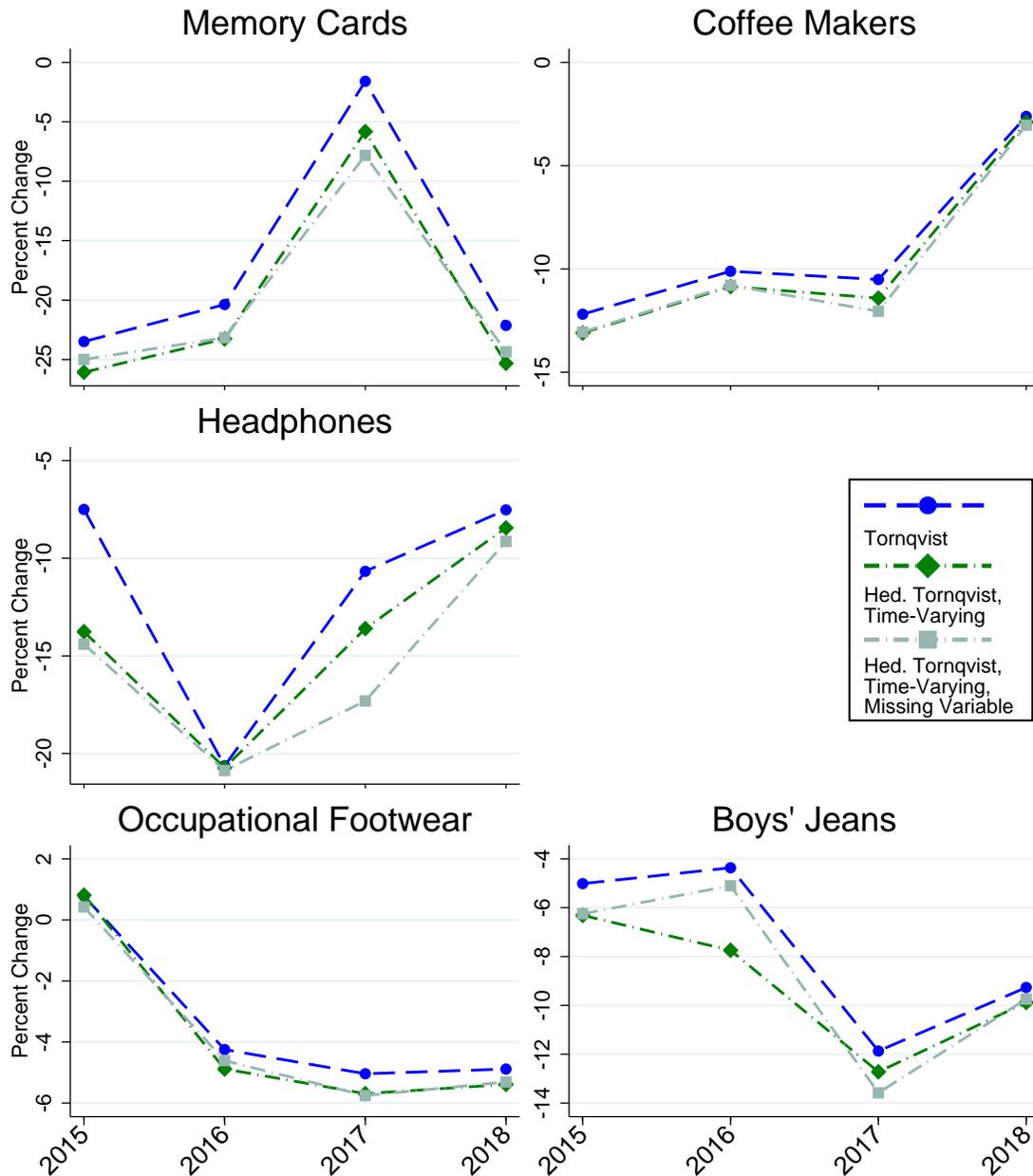
Notes: Unweighted average market share and prices relative to their value in the period of their initial entry. Entry occurs in period 0. All series are smoothed with a quartic spline. Data comes from the NPD Group.

Figure 3.2: Hedonic Specifications: Fixed vs. Time-Varying Unobservables



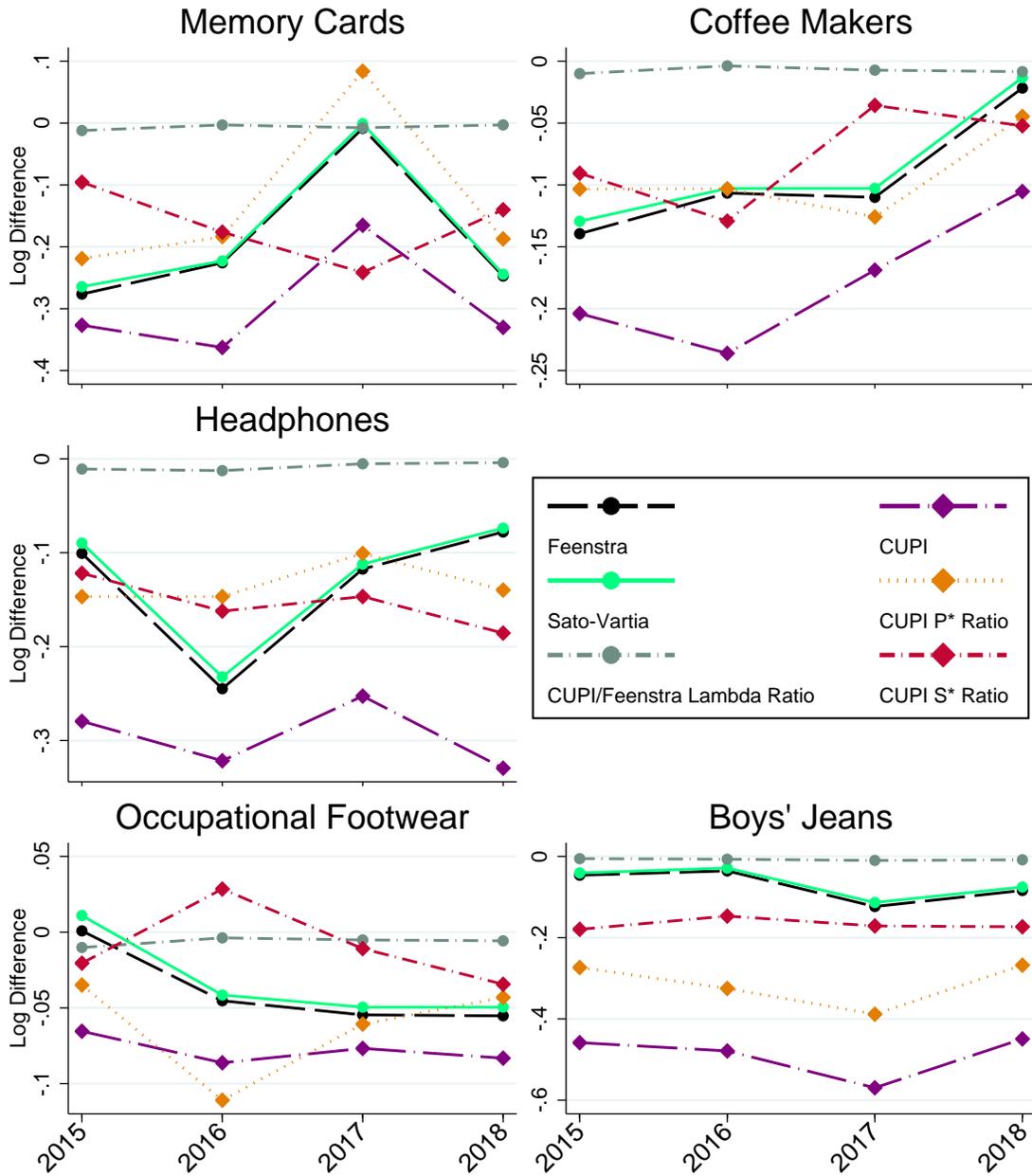
Notes: Values are percent change on a q4-to-q4 basis, aggregated from chained quarterly price indices. The time-dummy Tornqvist index uses adjacent period estimation with Tornqvist market share weights. The fixed unobservables model estimates hedonic models of log change in price using WLS and average quantity-share weights. The time-varying unobservables model adds lagged hedonic level residuals to the log-difference specification. Data comes from the NPD Group.

Figure 3.3: Hedonic Specifications: Evaluating Time-Varying Unobservable Specification



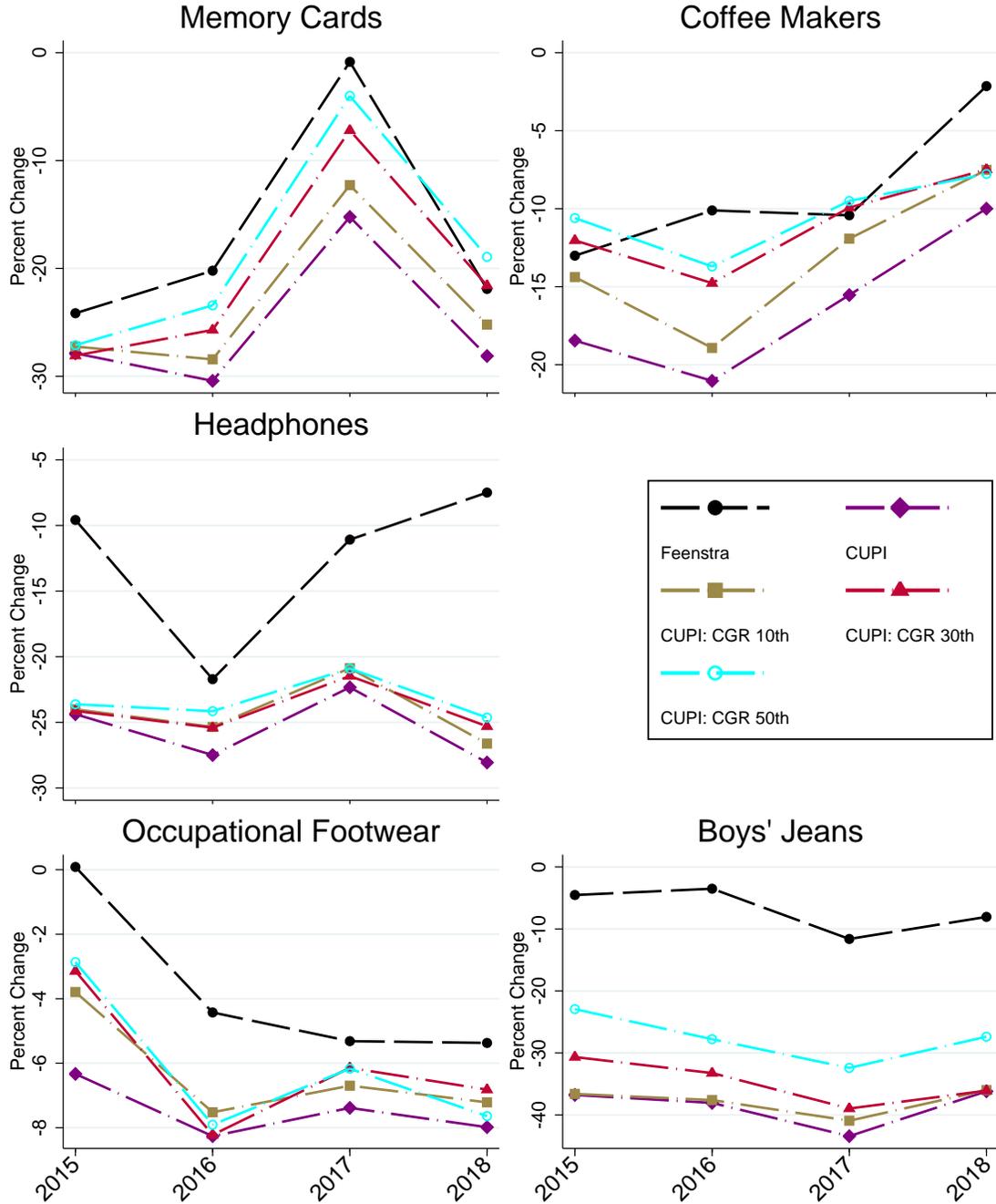
Notes: Values are percent change on a q4-to-q4 basis, aggregated from chained quarterly price indices. The time-varying unobservable model estimates hedonic models of log change in price using WLS and average quantity-share weights, including lagged hedonic level residuals. The “Missing Variable” series displays full imputation hedonic Tornqvist indices estimated using the time-varying unobservables approach, omitting key variables from the estimation. Data comes from the NPD Group.

Figure 3.4: Components of Feenstra and UPI



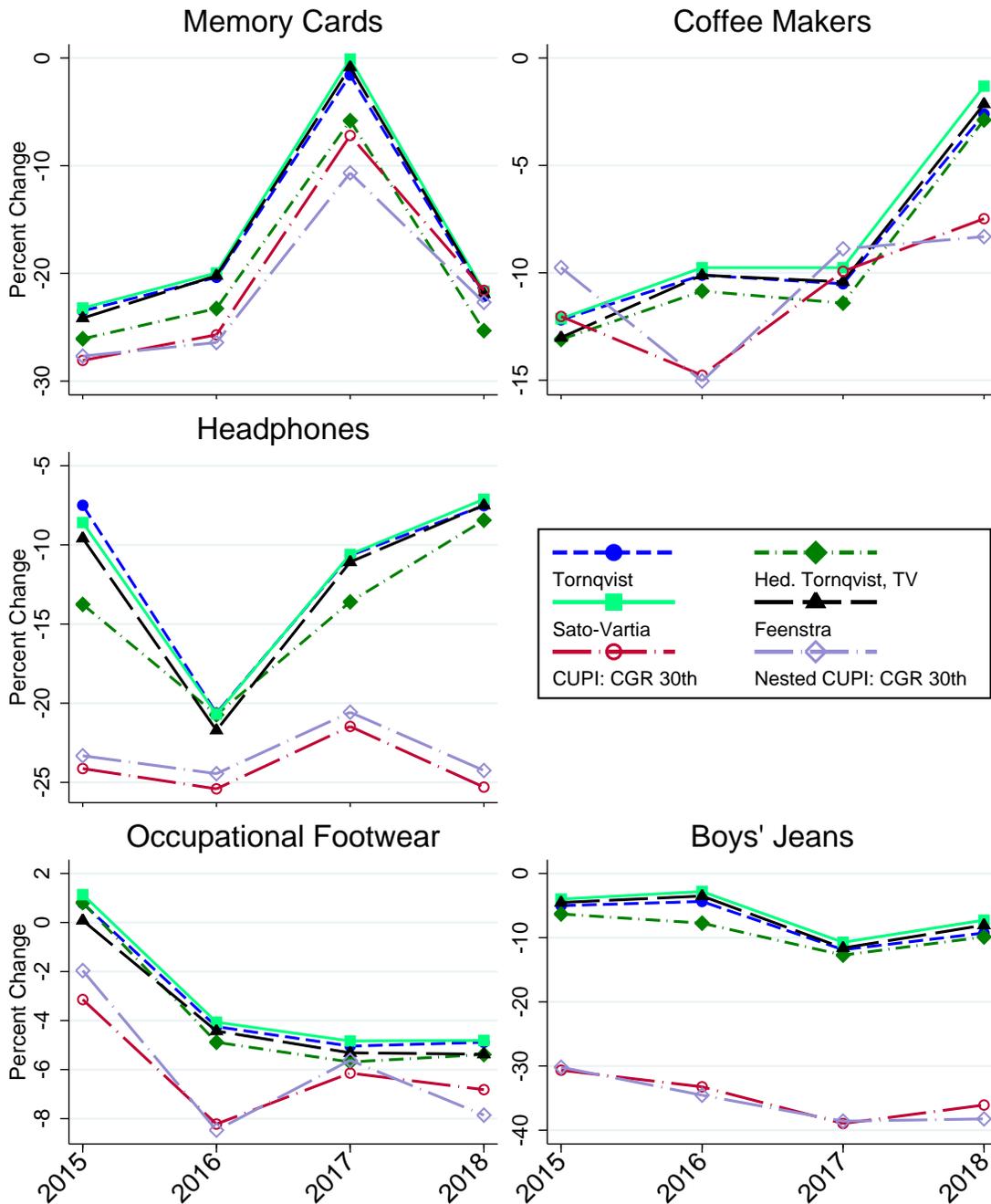
Notes: Values are log differences on a q4-to-q4 basis, aggregated from chained quarterly price indices. Units are reported in log-differences to allow for an additive decomposition of price indices. The Feenstra index is the sum of the Sato-Vartia and CUPI/Feenstra Lambda Ratio. The CUPI is the sum of the Lambda ratio, P*-ratio, and S*-ratio. Data comes from the NPD Group.

Figure 3.5: CUPI: Alternative Common Goods Rules (CGRs)



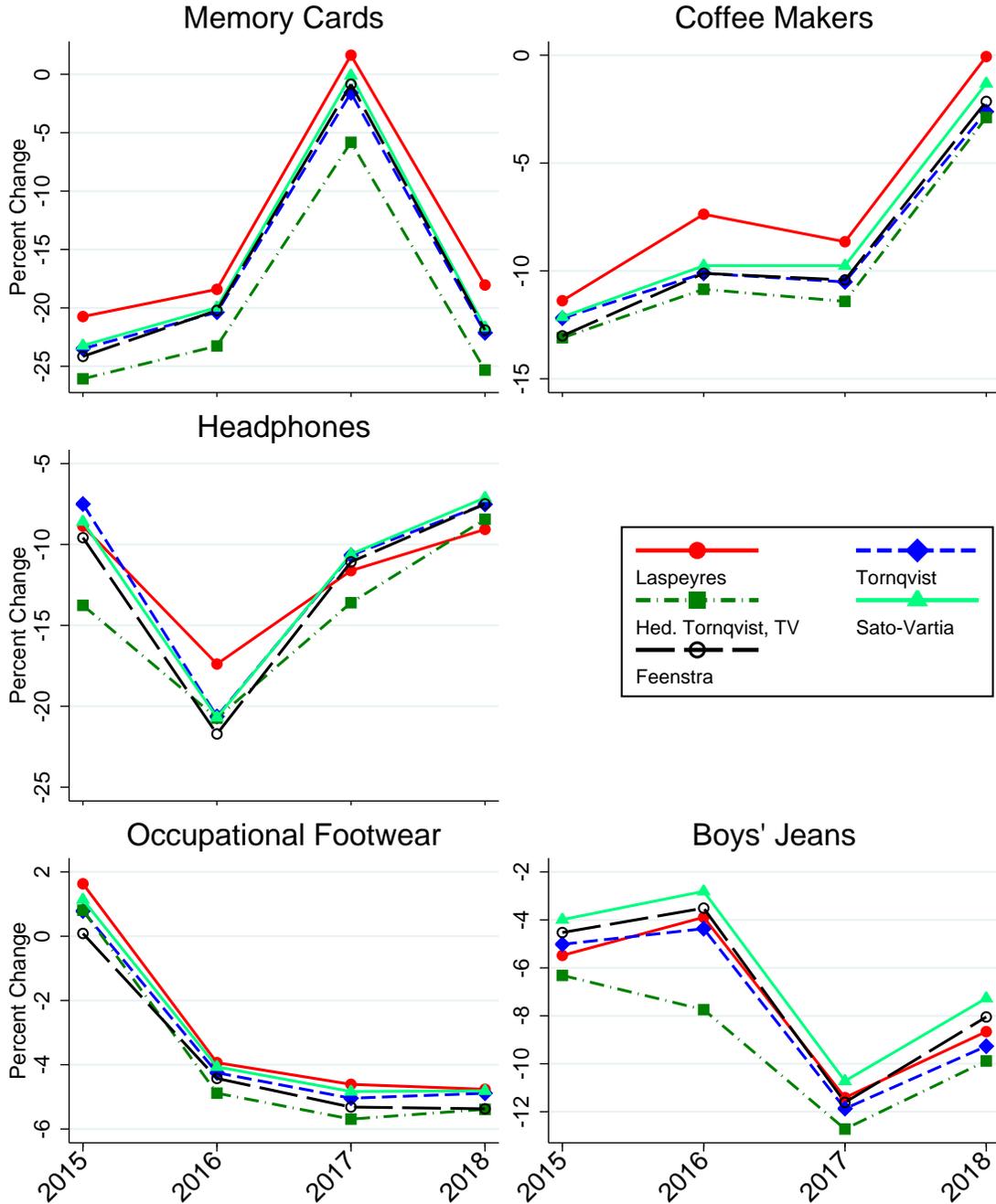
Notes: Values are percent change on a q4-to-q4 basis, aggregated from chained quarterly price indices. Common goods market share rules for the CUPI exclude from the group of common goods those products with market shares below the noted percentile in both periods. The Feenstra-adjusted Sato-Vartia index is included for reference. Data comes from the NPD Group.

Figure 3.6: Comparison of Main Price Index Specifications



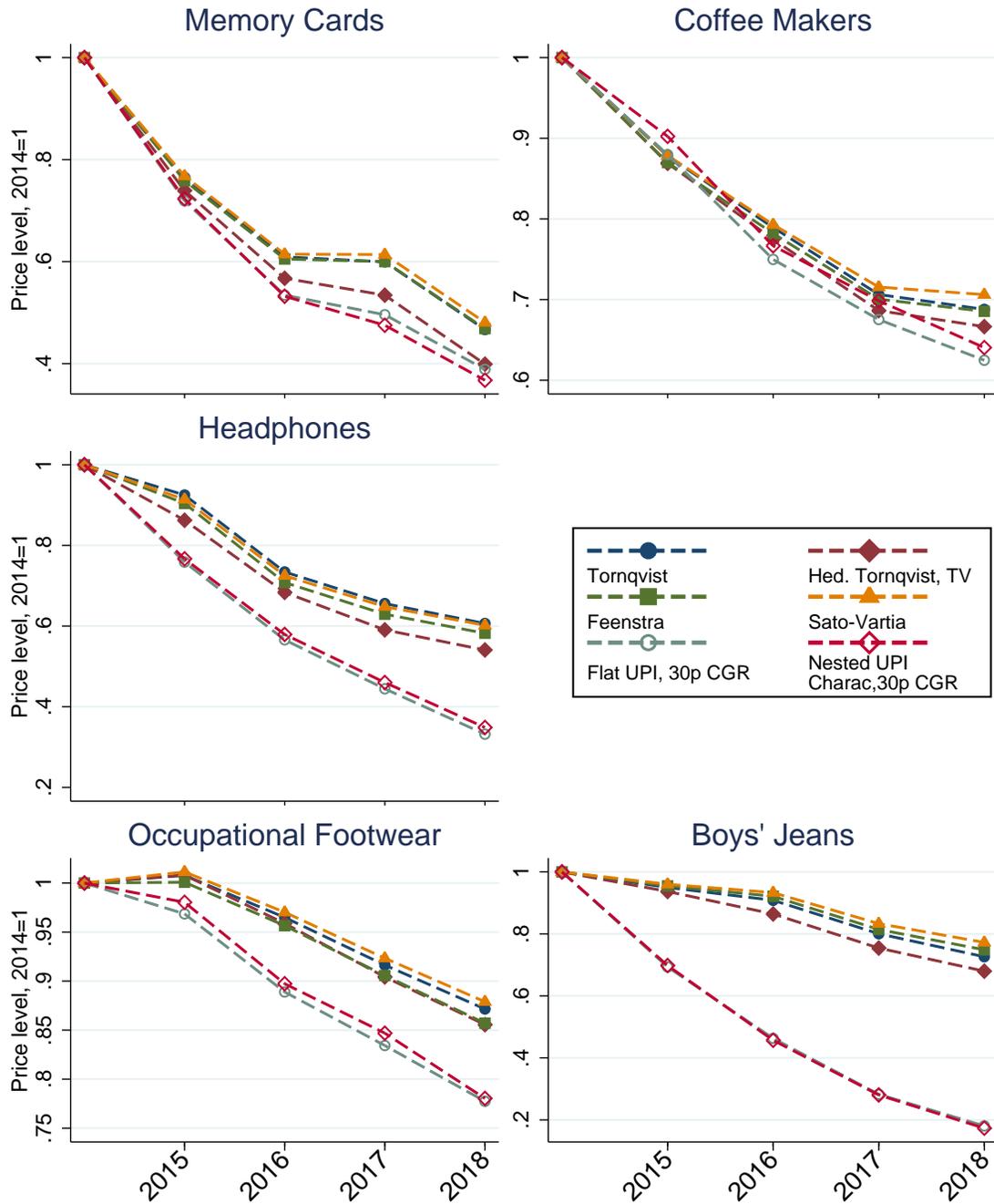
Notes: Values are percent change on a q4-to-q4 basis, aggregated from chained quarterly price indices. The hedonic time-varying unobservables model is estimated over log price differences using WLS and with weights that are average quantity-shares in adjacent periods. Data comes from the NPD Group. The Nested CUPI uses within-product-group nests based on observable characteristics. For the Nested CUPI, the 30th-percentile market share common goods rule is applied within nests.

Figure 3.7: Main Price Index Specifications, Without CUPI



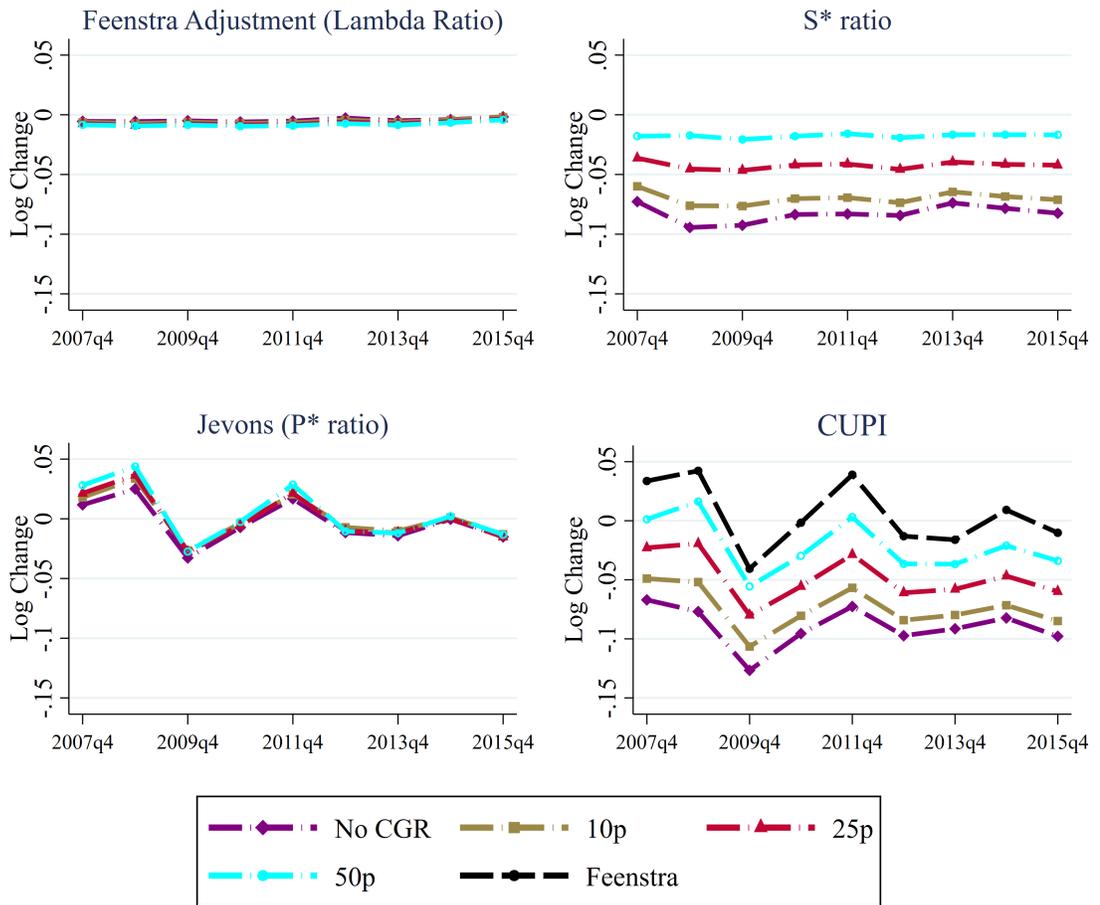
Notes: Values are percent change on a q4-to-q4 basis, aggregated from chained quarterly price indices. The Laspeyres series reports a geometric mean Laspeyres index. The hedonic time-varying unobservables model is estimated over log price differences using WLS and with weights that are average quantity-shares in adjacent periods. Data comes from the NPD Group.

Figure 3.8: Main Price Index Specifications: Cumulative Price Level Changes



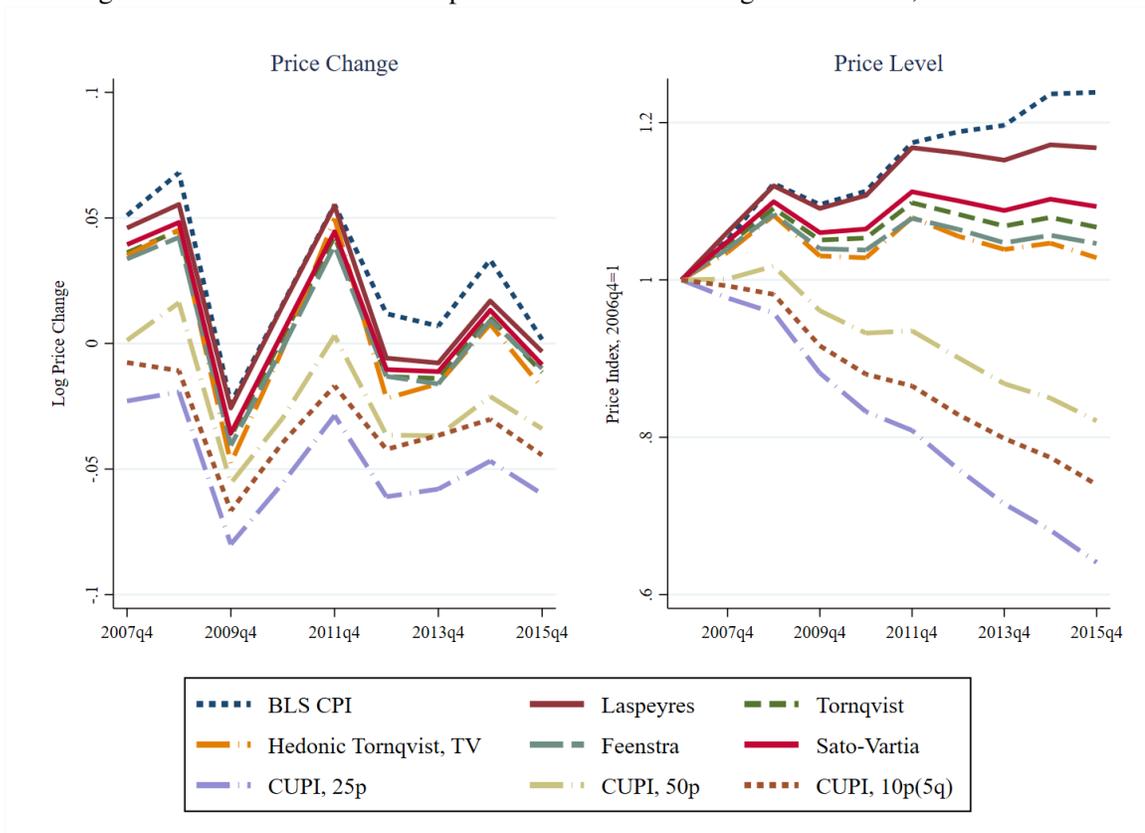
Notes: Values are cumulative changes relative to the 2014 price level, with 2014 price level set to 1. The hedonic time-varying unobservables model is estimated over log price differences using WLS and with weights that are average quantity-shares in adjacent periods. Data comes from the NPD Group.

Figure 3.9: CUPI and Its Components with Alternative CGRs, Nielsen Food



Notes: The figure shows Nielsen Retail Scanner data for food product groups. Each plot shows log changes from the fourth quarter of the previous year to the fourth quarter of the labeled year. The values are cumulative changes from chained quarterly indices. The Feenstra adjustment and S^* ratio panels show the adjustments scaled by $\frac{1}{\sigma-1}$ for each product group, so that the sum of those two components and the Jevons index (P^* ratio) equals the CUPI.

Figure 3.10: Main Price Index Specifications: Price Changes and Levels, Nielsen Food



Notes: The figure shows Nielsen Retail Scanner data for food product groups. Price changes show annual log differences from the fourth quarter of the previous year to the fourth quarter of the labeled year. The values are cumulative changes from chained quarterly indices. The price levels chained quarterly values of each price index in the fourth quarter of each year, with the price level in the fourth quarter of 2006 normalized to one for each index. The Laspeyres index is geometric. “CUPI, 25p” and “CUPI, 50p” use 25th-percentile and 50th-percentile quantity shares, respectively. “CUPI, 10p (5q)” uses a 10th-percentile quantity share threshold computed over quarters t and $t - 4$.

3.6 Appendix to Chapter 3

3.7 Additional Price Index Details

3.7.1 Traditional Price Indices

Generalizing equation (3.1), we can define every traditional geometric price index as a weighted average of log price changes. Specifically the log geometric price index, $\ln \Psi_t^G$, is given by

$$\ln \Psi_t^G \equiv \sum_{k \in \mathbb{C}_{t-1,t}} w_{kt} \ln \frac{p_{kt}}{p_{kt-1}},$$

where w_{kt} is a weight assigned to product k . The choice of weights determines the index. The Laspeyres index uses lagged expenditure shares as weights ($w_{kt} = s_{kt-1}^*$), the Paasche index uses current expenditure shares ($w_{kt} = s_{kt}^*$), and as seen in equation (3.1), the superlative Tornqvist index uses average expenditure shares ($w_{kt} = \frac{s_{kt-1}^* + s_{kt}^*}{2}$). Hence, the Tornqvist lies between the geometric Paasche and geometric Laspeyres. [Diewert \(2021\)](#) shows that for price indices at a detailed level of aggregation (so that the goods are sufficiently close substitutes), the “standard” ordering occurs - geometric Paasche < Tornqvist < geometric Laspeyres. The Sato-Vartia index uses logarithmic mean expenditure shares, as defined in equation (3.8).

Traditional price indices have a theory-free interpretation as weighted-average changes in product prices. While this statistical interpretation is valuable on its own, there is also an economic interpretation of these indices dating back to the seminal work of A.A. Konus ([Konüs, 1939](#); [Schultz, 1939](#)). The arithmetic Laspeyres and Paasche indices provide upper and lower bounds, respectively, on the exact change in the cost of living between two periods in the absence of product turnover and associated quality change.³⁴ Superlative indices, including the Fisher and Tornqvist, have more desirable theoretical properties:

³⁴In the case of strictly normal goods, the arithmetic Paasche is a lower bound of the equivalent variation, and the arithmetic Laspeyres is an upper bound to the compensated variation, so we have that Paasche \leq EV \leq CV \leq Laspeyres. Paasche < Laspeyres typically holds in the data, and will be the case when substitution is, on net, away from goods that have the highest change in price and towards those with the lowest.

they are the change in the unit expenditure function (i.e., the exact price index) that is the second-order approximation for a wide class of utility functions in the absence of product turnover and taste shocks (Diewert, 1978).³⁵

The traditional superlative price indices require both the *price* and *sales or expenditure share* of each good in either one or both time periods to calculate weighted price changes. As noted in the main text, however, in current practice, statistical agencies' data often does not permit the calculation of such weights. This practical limitation motivates the frequent use of the Laspeyres index in official statistics (e.g. for the CPI), which is subject to potentially large substitution bias relative to the superlative indices.

3.7.2 Hedonic Estimation Details

We estimate hedonic imputation models in both log-levels and log-differences as well as using the time dummy method. For the hedonic imputation models, we also consider alternative weighting approaches.

For the time dummy method, we specify the hedonic regression equation (3.5) using the same vector of characteristics Z_k in each pair of adjacent periods. Occasionally, new features are introduced to the data. In pairs of adjacent periods entirely prior to the introduction of a new characteristic, it will be omitted from the regression because of collinearity with the intercept term. In pairs of adjacent periods in which the new feature is absent during period $t-1$ and present during period t , the feature will be included in the estimated regression. Symmetric arguments apply for characteristics that exit.

Intuitively, the period- t fixed effect δ_t reflects the difference in average price of a “generic” good between $t - 1$ and t because the contributions of all of the product characteristics have been partialled out. The hedonic time dummy specification includes goods entering in period t and exiting after period $t-1$ through its use of the Tornqvist weights, which are average market shares between the two periods. Nonetheless, a limitation of the time

³⁵A longstanding question in the literature concerned whether the Sato-Vartia index is superlative, until Barnett and Choi (2008) demonstrated that it is. Like the Tornqvist, the Sato-Vartia index is also an expenditure-weighted average of log price changes; it differs from the Tornqvist by using the logarithmic mean of period $t-1$ and period t expenditure shares instead of the arithmetic mean.

dummy method relative to the TV approach is that the former does not account for unobservable product characteristics. Another issue emphasized by [Pakes \(2003\)](#) and [Diewert et al. \(2008\)](#) is that this method imposes constant coefficients on characteristics in adjacent periods, a restriction that is often rejected by the data.

Turning to the hedonic imputation indices, [Figure 3.A.1](#) presents results comparing the log-level relative to log-difference specifications. The log-level specification yields more erratic patterns than the log-difference specifications. We use quantity weights in the results presented in [Figure 3.A.1](#). For single-period log-level estimation, we use contemporaneous quantity shares. Intuitively, this specification only uses information from the current period to produce hedonic estimates. For estimation of the specifications proposed by [Erickson and Pakes \(2011\)](#), in which the dependent variable is the change in log prices, we use weights that are the average of the quantity shares in the previous and current periods. The results using the EP method presented in the main text take the same approach.

The log-level specifications are sensitive to omitted unobservable characteristics. To illustrate this point, [Figure 3.A.2](#) presents a version of [Figure 3.3](#) that shows the sensitivity of the levels specification to intentionally omitted key observable characteristics. Unlike the TV approach, the log-levels specification is very sensitive to omitting these observable characteristics.

The results presented thus far use quantity-share weights in our hedonic specifications following [Bajari et al. \(2021\)](#). As noted in the text, motivation for using quantity weights in estimation using unit prices is consistent with [Broda and Weinstein \(2010\)](#) and [Redding and Weinstein \(2020\)](#). The argument is that unit prices based on a large number of purchases are better measured than those based on a small number of purchases.

[Diewert \(2019\)](#) favors using expenditure-weights in hedonic estimation based on the argument that this provides more weight on items with more economic importance. He also notes that this approach facilitates comparisons of the time dummy and full imputation hedonic approaches. He acknowledges, however, that his preference for expenditure weighting is based more on index-number issues than econometric issues (footnote 14, p. 6 of

Diewert, 2019).³⁶

We report sensitivity of results using expenditure-weights in the regression-based TV results for NPD in Table 3.A.3. We present a large number of statistics including comparisons with related demand-based approaches to quality adjustment in Table . Given this, we focus on the results that incorporate chain drift using the GEKS-Lite procedure of Section 3.4.3. Such results reflect the many different issues including chain drift relevant for comparing alternatives. We report means and standard deviations of annual chained price indices as well as correlations. We included for comparison purposes the traditional Tornqvist and Sato-Vartia indices, the Feenstra index, and the hedonic Tornqvist using the TV approach estimated using both quantity weights and expenditure weights.

We find that using the expenditure-weighted and quantity-weighted approaches yield broadly similar results. For four of the five product groups, the quantity-weighted and expenditure-weighted TV price indices yield lower rates of annual price inflation than the traditional Tornqvist. The exception is headphones, for which the expenditure-weighted TV mean is slightly above the traditional Tornqvist. We also report and compare the hedonic Tornqvist indices' differences versus the traditional indices with the differences between the Sato-Vartia and Feenstra indices. For the latter, all five product groups exhibit a lower rate of inflation for Feenstra than Sato-Vartia. The patterns of these differences are more similar to the Tornqvist and Hedonic, TV using the quantity weighting. As discussed in the main text, we regard the demand-based quality approaches as a relevant benchmark to compare with the hedonic-based indices.³⁷

We also find that the quantity-weighted and expenditure-weighted indices are very highly correlated with each other and have similar variation as measured by the standard deviation. The traditional Tornqvist tends to have slightly lower correlations with the TV-based indices (especially for expenditure-weighted TV vs. traditional Tornqvist for Coffeemak-

³⁶Gorajek (2022) highlights that using expenditure weights yields potential issues with consistency of the estimation given that the expenditure shares are direct functions of the dependent variable (prices). He proposes an alternative transformation of the dependent variable to address these issues.

³⁷One note of caution about making such comparisons is that in principle, the Feenstra (1994) adjustment to the Sato-Vartia index captures pure love of variety effects in addition to quality improvement via entry and exit. The hedonic indices do not feature a direct role for gains from increasing product variety, although if increasing variety brings about lower prices or higher quality, it will affect the hedonic price indices indirectly.

ers). The Feenstra and TV-based indices have high correlations with the exception of coffee makers, for which the correlation is especially low using expenditure-weights.

We report goodness of fit statistics for alternative specifications in Table 3.A.4. As expected, the log-level estimation models account for a large share of variation in product price levels, as measured by R-squared. This high explanatory power reflects the fraction of the cross-sectional variation in prices accounted for by the observable characteristics. Those same models account for a small fraction of the variation in price relatives as can be seen in the second column. The EP methods (EP1 is first differences and EP2 is the TV approach) yield much higher R-squared values for the price relatives. The expenditure-weighted and quantity-weighted specifications yield similar R-squared values, although for occupational footwear and boys' jeans they are lower using expenditure-weighted estimation.

3.8 Using the Nielsen Data

3.8.1 Nielsen Data Preparation

The Nielsen RMS data consists of more than 100 billion unique observations at the week-store-UPC level. We first aggregate the weekly data to the monthly frequency according to the NRF calendar and then aggregate the monthly data to quarterly. Following procedures used by [Hottman et al. \(2016\)](#) and [Redding and Weinstein \(2020\)](#), we drop outliers from the monthly data before aggregating to the data to quarterly frequency. Specifically, we drop observations with prices above 3 times or below one-third the module-level median for each UPC in a given month. We also drop product-month observations with quantities sold that are more than 24 times that product's median quantity sold per month. One feature of barcoded products is that goods of different sizes and packaging have different barcodes, even if the product contained in the packaging is the same. To ensure comparability between prices, we follow [Hottman et al. \(2016\)](#) and normalize UPC prices to the same units (e.g., ounces), utilizing the size and packaging information provided by Nielsen. Consistent with the literature, we winsorize monthly price changes at the top and bottom 1% of each product group. Tables 3.A.6 and 3.A.7 show our classification of the product groups in the Nielsen Retail Scanner Data into Food and Nonfood categories, respectively.

3.8.2 Comparisons of the Nielsen Data to Official Statistics

In this section, we compare patterns of sales and prices for the Nielsen Scanner and Consumer Panel with official statistics. Using the Economic Census data from 2012, we have calculated that the types of retailers that the Nielsen scanner data tracks have very high coverage of food items (about 90%). Moreover, using a back-of-the-envelope calculation based on Nielsen's coverage of different types of retailers, we estimate that Nielsen scanner data accounts for about 41% of total food sales in the U.S. In contrast, the data's coverage is meaningfully lower for several nonfood categories. The types of stores Nielsen tracks accounts for about 53% of small appliance sales. However, Nielsen's coverage of general merchandise stores is only 32%. Using our back-of-the-envelope calculation, these figures imply that the Nielsen scanner data accounts for only about 19% of total small appliance sales in the U.S. Coverage in other categories is substantially lower. For instance, we estimate that the Nielsen scanner data accounts for only about 5% of total sales of hardware and tools.

We have also compared patterns of total expenditures for harmonized categories from Nielsen and Personal Consumption Expenditures data (PCE) from the Bureau of Economic Analysis. We have constructed a concordance between Nielsen and PCE categories at a detailed level (e.g., Bakery) and for broader categories—Food and Nonfood. For prices, we thank the BLS for preparing CPI indices for the broader categories of food and nonfood in a harmonized fashion.³⁸

Figure 3.A.7 presents comparisons of nominal expenditures for the broad food and nonfood categories. It is drawn from our companion paper [Cafarella et al. \(2023\)](#) and reproduced here for convenience. For food, we find nominal sales for the Nielsen Scanner data tracks the PCE closely. The Nielsen Consumer Panel tracks the PCE reasonably well through 2012, but it rises less rapidly than either the Nielsen Scanner or PCE thereafter. For nonfood, both the Scanner and Consumer Panel exhibit less of an increase over time than the

³⁸For the broad food and nonfood comparisons with PCE we use a concordance of the 100 plus product groups in the Nielsen data with the PCE. When we examine more detailed categories in Figure 3.A.8 we use a concordance provided to us by BLS between PCE categories and the 1000 or so Nielsen product modules. We have found that at the aggregate food and nonfood levels using the product level concordance or product module concordance is not important.

PCE.³⁹

These patterns are consistent with the discussion in the main text that the Nielsen data's coverage of nonfood items has deteriorated over time. Figure 3.A.8 provides more guidance on this point by showing for detailed categories the Scanner data the ratio of the growth in nominal sales for the Nielsen Scanner for the period 2008:1 to 2015:4 relative to the growth in nominal sales for the PCE over the same period. It is also drawn from our companion paper [Cafarella et al. \(2023\)](#) and reproduced here for convenience. The upper panel shows results for food categories and the lower panel for nonfood categories.⁴⁰ The categories from left to right in each panel are ranked by expenditure shares. Many of the food product categories have ratios close to one. In contrast, the nonfood categories have ratios that are much more variable and also typically below one.⁴¹

Figure 3.A.9 presents the relationship between the BLS CPI and corresponding Laspeyres indices from the Nielsen Scanner and Consumer Panel data sets. We show both arithmetic and geometric Laspeyres. The CPI is a two stage index with a geometric unweighted index at the MSA level and arithmetic Laspeyres to the national level. For food, both the Nielsen Scanner and Consumer Panel Laspeyres indices are highly correlated with the CPI. In terms of inflation levels, however, the Nielsen Scanner more closely matches the CPI (especially for the arithmetic Laspeyres using the Scanner data). The correlations between Laspeyres indices for the nonfood product groups and the CPI are much weaker than for food (0.53 and 0.67 for the Scanner and Consumer Panel data sets, respectively, using the arithmetic Laspeyres). The average inflation level is closer to the CPI in the Scanner data than in the Consumer Panel.

We interpret these results as providing justification for our focus on food results using the Nielsen Scanner data in the main text. The results also support the view that the Nielsen Scanner data tracks the official statistics as well as, if not more closely than, the Nielsen

³⁹For our analysis of the Retail Scanner we use the NRF calendar, while for the Consumer Panel we use the regular calendar. This difference is not important for the patterns reported in this and the next sections. The NRF calendar is especially relevant at the monthly frequency.

⁴⁰Even though our concordance is at the product module level, the categories are more aggregated than Nielsen product groups.

⁴¹The results presented by detailed category are for the Nielsen Scanner data, which is the primary focus of our analysis. In unreported results, we find similar patterns for the Nielsen Consumer Panel data.

Consumer Panel.

3.8.3 Product Descriptions in the Nielsen Data

The product descriptions in the Nielsen data provided by the Kilts Center for Marketing at the University of Chicago are generally not coded to be human-intelligible. For instance, two product descriptions for soft drinks are ZR DT LN/LM CF NBP CT and NATURAL R CL NB 12P, while a product description for toilet paper is DR W 1P 308S TT 6PK. A human analyst could decipher portions of these descriptions: DT means “diet,” 12P means twelve pack, 1P means one ply, 308S means 308 sheets, etc. It would not be feasible for human analysts to encode such data at scale, however, and simple dictionaries would be fooled (e.g., the P-suffix means “pack” for soft drinks and “ply” for toilet paper).

An additional challenge in the Nielsen data is its sheer scale. The Retail Scanner data contains more than 100 product groups and over 1,000 product modules. It would be difficult for human analysts to specify sensible hedonic regression equations for so many product groups. It would be even more difficult to update those regression equations over time as product mixes and characteristics change.

3.8.4 Common Goods Rules – Consumer Panel and Retail Scanner

This section presents sensitivity results to alternative common good rule approaches for both the Nielsen Scanner and Nielsen Consumer Panel data sets. Using the scanner data, Figure 3.A.10 compares the results of imposing common goods rules using the 2-quarter horizon, as in the main text (i.e., using percentiles from sales pooled over the current and prior periods), vs. a 5-quarter horizon (i.e., computing percentiles for sales pooled over quarters t and $t - 4$).⁴² These alternative CGRs impose different duration-based restrictions on products to be included in the set of common goods. The 2-quarter horizon CGR requires goods to be present in periods t and $t-1$, while the 5-quarter horizon requires goods to be present in periods t and $t - 4$. The figure shows that the 5-quarter CGR using a 10th-percentile share threshold leads the CUPI to measure inflation between what is measured

⁴²In many of the figures of this appendix, we include the arithmetic Laspeyres as this facilitates comparison with Redding and Weinstein (2020). The prior section shows arithmetic and geometric Laspeyres yield similar patterns.

using 25th and 50th percentile thresholds using the 2-quarter horizon. The longer-horizon CGR puts additional weight on the goods that have been present in the marketplace for a longer time, which moves our approach in the direction of the duration-based CGR approach of [Redding and Weinstein \(2020\)](#).

Figure 3.A.11 shows the sensitivity of the CUPI to different CGRs using the Nielsen Consumer Panel for food. Here, we focus on 5-quarter horizon CGRs. While the results differ quantitatively, the same general pattern holds as in the Nielsen Scanner data, with the CUPI increasing in the percentile of the CGR.

To facilitate comparison of our results to [Redding and Weinstein \(2020\)](#), who report pooled results for food and nonfood product groups, Figure 3.A.12 shows various price indices calculated using all product groups in the Nielsen Consumer Panel data. The results are broadly consistent with [Redding and Weinstein \(2020\)](#). However, importantly our analysis focuses on chained quarterly annual indices while [Redding and Weinstein \(2020\)](#) focus on year-over-year indices for fourth quarters of each year. In Figure 3.A.13, we show we can closely mimic their results for the CUPI using a market share common goods rule at the 5th percentile if we calculate a Y-o-Y price index instead of the chained quarterly price indices that have been the focus of this paper. As we have noted in the preceding discussion, the use of a Y-o-Y index imparts a duration-based component to the CGR in addition to the expenditure share-based thresholds.⁴³

Figure 3.A.14 shows related indices, using the Nielsen Scanner data, pooling all product groups, and using various CGRs based on sales percentiles computed over the 5-quarter horizon.⁴⁴ These results are therefore suggestive of the results applying the empirical approach in [Redding and Weinstein \(2020\)](#) to the Scanner data would produce. The CUPI with no CGR suggests deflation of 10 percent or more per year. Even the CUPI with a

⁴³We note that we do not impose a CGR in computing the other price indices shown in Figure 3.A.12. In contrast, [Redding and Weinstein \(2020\)](#) apply the same common goods rule for all of the price indices they display. In unreported analysis, we have found that the Sato-Vartia and Feenstra are not very sensitive to the CGR. This inference is also evident in Figure 3.9 that shows that is sensitive to the CGR for Nielsen Food data. Because our objective is to compare demand-based indices with the hedonic indices, we aim to treat entry and exit symmetrically across these indices.

⁴⁴Figure 3.A.14 also displays the Bureau of Labor Statistics' Consumer Price Index for all of the product groups included in the Nielsen data as a point of reference.

25th-percentile cutoff rule shows persistent deflation in the Retail Scanner data; imposing a 50th-percentile CGR brings the CUPI closer in line with the Laspeyres index. The series labeled “CUPI, RW CP” shows results from applying the market share threshold in the 5th-percentile CGR from the Consumer Panel to the Scanner Panel data, rather than calculating a percentile-based threshold directly from the Scanner Panel data. Using the Consumer Panel share threshold for the CGR produces results similar to using the 50th-percentile CGR calculated directly in the Scanner Panel data.

The lower inflation rates the CUPI measures in the Nielsen Retail Scanner data relative to the Consumer Panel data highlight the scanner data’s large number of very low-market share products. This long tail disproportionately impacts the CUPI. In contrast, the Laypeyres and Feenstra indices are much more consistent between the Nielsen Consumer Panel and Nielsen Retail Scanner data.

Figure 3.A.15 displays for the nonfood product groups the analogous plots to Figure 3.10, which displays results for food product groups. For comparability purposes to the those in the main text, the CGR rules in this figure are based on sales percentiles over the 2-quarter horizon.⁴⁵

The main message from this analysis is that the CUPI is very sensitive to the specification of the CGR, both in the Nielsen Consumer Panel and in the Nielsen Scanner data. This sensitivity applies both to the market share threshold used and to the horizon over which the threshold is computed. Using the longer horizon market share threshold moves the CGR towards the Redding and Weinstein (2020) duration-based approach. It is worth reiterating that any duration based approach has greater data requirements for practical implementation.

3.8.5 Machine Learning and Hedonics

This appendix summarizes our procedure for incorporating machine learning into hedonic estimation. Our companion paper, Cafarella et al. (2023), provides further details.

Using machine learning (ML) methods to estimate hedonic price indices requires making

⁴⁵To be consistent with the results for food reported in the main text, Laspeyres is geometric in this figure.

several practical choices regarding the architecture of the ML system used for prediction and the conversion of those predictions into price indices. As discussed in the main text of this paper, our preferred approach to constructing hedonic price indices is the “time-varying unobservables” hedonic imputation approach of [Erickson and Pakes \(2011\)](#). The core of this method is to estimate price *levels* for each product in each period in a first step. In a second step, this approach estimates price *changes*, using the hedonic residual (or prediction error) from the first step as a predictor. This methodology allows the hedonic predictions partially to capture unobserved product characteristics’ influence on price changes.

In many ways, the “TV” approach of [Erickson and Pakes \(2011\)](#) can incorporate ML methods quite naturally. The key innovation is to use ML methods rather than standard regression techniques to estimate the hedonic functions for log price levels and changes in equations (3.2) and (3.4). Another important difference from the more standard econometric procedures we employ in the NPD data is that the Nielsen data available from the Kilts Center does not include pre-coded item-level product attributes. Attribute information is limited to short, non-standard text descriptions. We use deep neural networks to predict product prices and price changes from these product descriptions.

Several features of our methodology merit particular discussion. First, to convert text-based product descriptions into numerical characteristic representations, we use a hybrid feature encoding architecture that allows the system to incorporate “pre-trained” word embeddings (numerical representations) trained from an external corpus of text as well as specifically trained or “text-tailored” embeddings trained specifically on the product descriptions in the Nielsen Kilts Retail Scanner Data set. Second, our architecture does not predict prices or price changes directly, but rather predicts a set of probabilities that the price or price change lies in each of a set of price or price-change bins that partition the observed range. Third, the ML system minimizes the weighted cross-entropy loss function for the products’ true price and price change distributions in the hedonic estimation.⁴⁶ Both steps are weighted using products’ unit sales (quantities) shares in a product-group quarter. Fourth, because of the noise in the estimated probabilities, it may not be optimal to calculate price predictions as

⁴⁶In this application, the cross-entropy loss objective function is equivalent to maximizing the likelihood of assigning the highest probability to the correct bin.

the simple probability-weighted expected price. We use a receiver operating characteristic (ROC) curve procedure to determine the optimal number of bins to include in the price prediction.

[Cafarella et al. \(2023\)](#) explores the ML procedure’s performance as measured by the prediction “near accuracy” across every product group-quarter. We define the model’s near accuracy as the proportion of products for which it assigns the highest probability to the correct or an adjacent bin. The median in-sample near accuracy for food price change bins is well above 80%. The out-of-sample near accuracy for the median product group-quarter is nearly 60% for the food product groups. In other words, the median-performing model predicts the the correct bin or an adjacent bin more than 80% of the time. We view these model performances as remarkable: in the median product group-quarter, the system is able to closely predict a product’s price change over half the time based on the short, nonstandard product descriptions.

3.8.6 Assessing the National Market Assumption in the Nielsen Data

As noted in Section [3.4.4](#), our empirical implementation of the CUPI relies on the assumption of a unified national market. Because the CUPI is not consistent in aggregation and includes unweighted geometric mean terms, any failure of this assumption may affect the CUPI more than other indices. In this section, we assess the realism of the national market assumption in the Nielsen scanner data.

In Figure [3.A.6](#), we pool the Nielsen item-level data for food product groups at the weekly frequency from 2006–15. We then compute the market penetration of items in the pooled data both on an unweighted basis (i.e., all items get the same weight) and on a sales-weighted basis. Market penetration is defined as the share of Nielsen metro areas in which the item-level week is observed to have positive sales.

On an unweighted basis, the distribution is very skewed to the left, with most item-level week observations having very low market penetration. Almost all of the unweighted distribution has less than a 20 percent market penetration. In unreported results, we find that the mass of the unweighted distribution with the lowest market penetration reflects entering

and exiting goods. Even on a sales-weighted basis, only 15 percent of sales are for items with a truly national market, although much of the mass of the distribution has market penetration of over 80 percent of metro areas. These patterns raise questions about applying a national market based CES price index for most items. In Appendix 3.9.3, we show that the CUPI can be badly biased when the national market assumption fails, suggesting these patterns may be important for understanding the empirical behavior of the CUPI.

3.9 Additional Evidence on the Behavior of the CUPI

In this appendix, we examine the behavior of the CUPI in additional detail. We begin in Section 3.9.1 by examining empirically whether implementing a nested structure in the CUPI modifies its extremely negative inflation readings. We find that the nesting nesting approaches we have explored do not meaningfully modify the CUPI's measurement of inflation. We then examine analytically and via simulation studies whether time-varying product appeal shocks generate an expected bias in the Sato-Vartia index relative to the consumer's exact price index under CES preferences. In section 3.9.2, we examine the mathematical source of the taste shock bias highlighted by Redding and Weinstein (2020). We conclude that the presence of time-varying product appeal on its own will not generate an expected bias in the Sato-Vartia index. On the other hand, time trends in the *dispersion* of product appeal shocks do introduce an expected bias. We show that this explanation is empirically unlikely, however, to account for the extremely low inflation measured by the CUPI.

A natural question that arises from that conclusion is why the CUPI measures consistently lower inflation than the Sato-Vartia and Feenstra indices. In section 3.9.3, we present simulation evidence showing that geographical segmentation of entering goods and limited availability of existing goods can cause the CUPI to measure significantly lower inflation than is implied by the consumer's unit expenditure function. A common goods rule helps alleviate such biases. We believe these simulations point the way toward future research on the implementation of the CUPI.

3.9.1 Behavior of the CUPI with a Nested Preference Structure

We explore the issue of nesting in the CUPI using two methods that rely on the product attributes in the data to define a nested product substitution structure. First, we define nests within product groups with a heuristic-based approach. Using this approach, we assign products to subgroups based on a set of key variables that we as analysts hypothesize define market strata. Because this procedure is labor-intensive and relies on our subjective judgments regarding strata, we also construct alternative subgroups by allocating products to groups based on the deciles of their predicted price from a log-level hedonic model. Intuitively, in the first approach, we implicitly assume that substitutability is constant within market strata (for example, drip coffee makers versus espresso machines), while in the second approach we assume that price tiers (for example, low-end versus high-end coffee makers) define the substitution structure.

The nested approach requires estimation of elasticities of substitution for products within the same nest and across nests. We follow the approach of [Hottman et al. \(2016\)](#) to estimate within- and between-nest elasticities for each product group. The within-group estimation uses a modified [Feenstra \(1994\)](#) estimator that double-differences market shares and prices with respect to time and a time-varying nest-level mean.⁴⁷ The between-nest estimator of the elasticity of substitution uses an instrumental variable (IV) approach building on [Hottman et al. \(2016\)](#).⁴⁸

Table 3.A.5 reports the estimated elasticities for the nested specifications. The results are broadly similar across the two nested approaches. As expected, the within-nest elasticities are estimated to be larger than the between-nest elasticities.

⁴⁷The identifying assumption of the [Feenstra \(1994\)](#) estimator is that supply and demand shocks are orthogonal when sales growth and price growth are differenced with respect to a time-varying mean. The [Hottman et al. \(2016\)](#) assumption is arguably more natural, as differencing with respect to a within-nest mean more plausibly identifies orthogonal supply and demand shocks.

⁴⁸We follow [Hottman et al. \(2016\)](#) by specifying the between-group relationship between the nest-level price index and expenditure share. The former is endogenous, and [Hottman et al. \(2016\)](#) overcome this by using variation in the nest-level price index caused by changes in within-nest expenditure share dispersion. We innovate on the procedure of [Hottman et al. \(2016\)](#) by using the S^* ratio (i.e., changes in common goods expenditure share dispersion) from the within-nest CUPI as the instrument, which removes changes in expenditure-share dispersion induced by product turnover. The identifying assumption is that within-nest demand shocks are uncorrelated with between-nest demand shocks. This innovation integrates the insights of [Hottman et al. \(2016\)](#) with those of [Redding and Weinstein \(2020\)](#).

In principle, these within-nest vs. between-nest elasticity estimates could produce significantly different results for the Feenstra index and the CUPI, but in our application the differences are modest. Figure 3.A.4 plots nested versions of the CUPI using our two nesting strategies alongside un-nested versions of the CUPI and Feenstra index. Both versions of the CUPI are implemented using a 30th-percentile CGR, applied at the within-nest level in the nested version.⁴⁹ The alternative nesting approaches yield similar results, with the nested CUPI tending to show slightly less deflation than the un-nested (or “flat”) CUPI. In unreported results, we find that the relationship between the nested and flat CUPIs is robust to using alternative CGR cutoffs.

3.9.2 Analytical Characterization of the Taste Shock Bias

We consider a representative consumer with CES preferences. For simplicity, in this subsection we examine a market with no product turnover, and we assume the consumer has non-nested preferences over the set of available products. Let N denote the number of products present in each period and P_t denote the unit expenditure function in period t . Redding and Weinstein (2020) show that, in the presence of product appeal shocks, the change in the log Sato-Vartia price index equals the change in the log unit expenditure function plus an additional term

$$\ln \Phi_t^{SV} = \ln \frac{P_t}{P_{t-1}} + \left[\sum_k \omega_{kt} \ln \left(\frac{\varphi_{kt}}{\varphi_{kt-1}} \right) \right], \quad (3.9.1)$$

where ω_{kt} are the Sato-Vartia weights defined by

$$\omega_{kt} = \frac{\frac{s_{kt} - s_{kt-1}}{\ln(s_{kt}) - \ln(s_{kt-1})}}{\sum_k \frac{s_{kt} - s_{kt-1}}{\ln(s_{kt}) - \ln(s_{kt-1})}}.$$

Redding and Weinstein (2020) label the term in the square brackets of equation (3.9.1) the “taste shock bias,” as it represents the difference between the Sato-Vartia index and the true change in the cost of living index. It is easy to see that when product appeal is constant over time, so that $\varphi_{kt} = \varphi_{kt-1}$ for every product k , the taste shock bias term will be zero and

⁴⁹Nests are weighted by the number of products to adjust for differential product group sizes.

the Sato-Vartia index will exactly recover the true change in the cost of living. [Redding and Weinstein \(2020\)](#) argue that when product appeal is time varying, however, the taste shock bias term will be positive in expectation, so that the Sato-Vartia index will tend to overstate the true rate of inflation.

The expected taste shock bias can be written as

$$\begin{aligned}\mathbb{E} \left[\ln \Phi_t^{SV} - \ln \frac{P_t}{P_{t-1}} \right] &= N\mathbb{E} \left[\omega_{kt} \ln \left(\frac{\varphi_{kt}}{\varphi_{kt-1}} \right) \right] \\ &= NCov \left[\omega_{kt}, \ln \left(\frac{\varphi_{kt}}{\varphi_{kt-1}} \right) \right] + N\mathbb{E} \left[\ln \left(\frac{\varphi_{kt}}{\varphi_{kt-1}} \right) \right].\end{aligned}\quad (3.9.2)$$

The second term in equation (3.9.2) will be zero due to the normalization. [Redding and Weinstein \(2020\)](#) note, however, that the Sato-Vartia weights ω_{kt} are an increasing function of the appeal parameters φ_{kt} , so

$$\frac{\partial \omega_{kt}}{\partial \varphi_{kt}} = \frac{\partial \omega_{kt}}{\partial s_{kt}} \frac{\partial s_{kt}}{\partial \varphi_{kt}} > 0 \implies Cov \left[\omega_{kt}, \ln \left(\frac{\varphi_{kt}}{\varphi_{kt-1}} \right) \right] > 0. \quad (3.9.3)$$

Other factors equal, consumers will devote a greater share of expenditure to goods that experience favorable appeal shocks. In isolation, that tendency would lead the Sato-Vartia taste-shock bias in equation (3.9.2) to be positive. As [Redding and Weinstein \(2020\)](#) argue in their abstract:

In the presence of relative taste shocks, the Sato-Vartia price index is upward biased because an increase in the relative consumer taste for a variety lowers its taste-adjusted price and raises its expenditure share. By failing to allow for this association, the Sato-Vartia index underweights drops in taste-adjusted prices and overweights increases in taste-adjusted prices, leading to what we call a “taste-shock bias.”

We believe that this intuition, while correct on its own, is also incomplete: there is a symmetrical and offsetting tendency for appeal shocks to induce a downward bias in the Sato-Vartia index when the appeal parameters φ_k are independently and identically distributed across periods $t-1$ and t . The offsetting bias comes from the fact that the Sato-Vartia weights

ω_{kt} are also an increasing function of the *previous* period's appeal parameters φ_{kt-1} , which enter the second term in the covariance, $\ln\left(\frac{\varphi_{kt}}{\varphi_{kt-1}}\right)$, in the opposite direction from the current period's appeal parameters. Hence,

$$\frac{\partial \omega_{kt}}{\partial \varphi_{kt-1}} = \frac{\partial \omega_{kt}}{\partial s_{kt-1}} \frac{\partial s_{kt-1}}{\partial \varphi_{kt-1}} > 0 \implies \text{Cov}\left[\omega_{kt}, \ln\left(\frac{\varphi_{kt}}{\varphi_{kt-1}}\right)\right] < 0. \quad (3.9.4)$$

This offsetting tendency would lead the Sato-Vartia taste-shock bias to be negative in isolation. The upward and downward biases will offset each other in expectation when the appeal parameters are identically distributed across periods $t-1$ and t , so the Sato-Vartia index will not exhibit a generic taste-shock bias under those assumptions.

Nonetheless, if the assumption of idiosyncratically and identically distributed appeal parameters does not hold precisely, for instance, because the dispersion of product appeal changes over time, the Sato-Vartia price index may exhibit a taste-shock bias. In particular, as noted by [Redding and Weinstein \(2020\)](#), increasing dispersion in product appeal will induce an upward bias in the Sato-Vartia index, while the CUPI will remain unbiased.

We examine the empirical relevance of this explanation for the empirical gap in [Figure 3.A.5](#). The figure plots the measured dispersion (standard deviation) in normalized product appeal for each of the NPD product groups. We find evidence of rising relative product appeal dispersion for memory cards and headphones. This increase is more pronounced without a common goods rule. For other goods, we observe a nonmonotonic change in dispersion even where the S^* ratio is highly negative (see, e.g., boys' jeans in [Figure 3.4](#)). In other words, the CUPI's ubiquitous finding of rapid deflation without using a common goods rule is not readily justified by the observed patterns of dispersion in product appeal.

3.9.3 Simulation Evidence on the Behavior of the CES Exact Price Indices

Simulation Model Environment

We base our simulations on the general equilibrium environment of [Hottman et al. \(2016\)](#).⁵⁰ A set of firms, indexed by f , each produces multiple products, indexed by u . Consumers have nested preferences, with preferences over the total output of each firm in the upper-level nest and preferences over the individual products supplied by each firm in the lower-level nests.

The bottom-level CES consumption index over the products supplied by firm f , C_{ft}^F , is given by

$$C_{ft}^F = \left[\sum_{u \in \Omega_{ft}^U} (\varphi_{ut} C_{ut})^{\frac{\sigma^U - 1}{\sigma^U}} \right]^{\frac{\sigma^U}{\sigma^U - 1}}, \quad (3.9.5)$$

where C_{ut} represents the quantity consumed of product u in period t , φ_{ut} is a product-level appeal shifter for product u , Ω_{ft}^U is the set of products supplied by firm f in period t , and σ^U is the elasticity of substitution among the products supplied by a firm.

The consumer's utility from consuming the output supplied by all firms, U_t , is given by

$$U_t = \left[\sum_{f \in \Omega_t^F} (\varphi_{ft} C_{ft}^F)^{\frac{\sigma^F - 1}{\sigma^F}} \right]^{\frac{\sigma^F}{\sigma^F - 1}}, \quad (3.9.6)$$

where C_{ft}^F is the firm-level consumption aggregate defined in equation 3.9.5, φ_{ft} is a firm-level appeal shifter for firm f , Ω_t^F is the set of firms supplying products in the marketplace in period t , and σ^F is the elasticity of substitution across firm-level consumption aggregates.

It is necessary to provide a normalization for the product-level and firm-level appeal shifters

⁵⁰[Hottman et al. \(2016\)](#) consider consumers with Cobb-Douglas preferences over a number of different product groups and constant elasticity of substitution (CES) preferences within each product group. For simplicity, we restrict our attention to consumers with preferences over products within a single group.

φ_{ut} and φ_{ft} . We follow [Redding and Weinstein \(2020\)](#) in assuming that both product-level and firm-level appeal shifters for continuing products have an average log change of zero in every period. That normalization allows for the possibility that entering or exiting products have higher or lower average appeal levels than continuing products.

The supply side of the market is populated by a set of firms that produce output using a composite input factor that serves as the economy's numeraire good. Firms' cost functions are assumed to be additively separable across products supplied. The total variable cost of producing Y_{ut}^U units of product u at time t , A_{ut} , is given by

$$A_{ut} (Y_{ut}^U) = a_{ut} (Y_{ut}^U)^{1+\delta}, \quad (3.9.7)$$

where a_{ut} is a marginal cost shifter of producing product u at time t , and δ is the elasticity of marginal costs with respect to output. We assume that product entry and exit is exogenous in our simulations.

Firms choose prices under Bertrand competition. Each firm's decisions affect other firms' decisions only through their effects on the economywide price index. In equilibrium, firms choose product prices to maximize profits and consumers choose quantities demanded of each product. We will generally assume that the market clears so that $C_{ut} = Y_{ut}^U$ for every product u and time t . Certain simulations will feature market imperfections that prevent this market-clearing condition.

[Hottman et al. \(2016\)](#) derive analytical formulas for consumers' product demands, firms' pricing rules, and firm-level and aggregate price indices in this environment, and provide computer code to solve for the market-clearing general equilibrium numerically. They also characterize the economics of the market environment in depth. We build our numerical simulations on the code provided by [Hottman et al. \(2016\)](#), so our environment will parallel theirs except for the differences that we highlight to explore the behavior of the CES exact price indices in various market environments.

Each simulation contains 50 firms and lasts for 40 periods.⁵¹ Unless otherwise noted, each

⁵¹We initialize all stationary variables by drawing from their steady-state distributions, so the simulations

firm sells 50 products in each period. 100 Monte Carlo simulations were run for each set of model parameters considered. To abstract from issues of within-firm vs. between-firm substitution, we set the elasticity of substitution between a firm's individual products σ^U equal to the elasticity of substitution across firms' composite output σ^F .⁵² We choose a value of 5 for both elasticities, between the values of σ^U and σ^F in [Hottman et al. \(2016\)](#) of 7 and 4, respectively. We set the elasticity of marginal costs with respect to quantity supplied δ to 0.15, consistent with the Monte Carlo simulations in [Hottman et al. \(2016\)](#).

Each period, the log product appeal shifters are drawn from normal distributions with product-specific means and variances. The product-specific means are drawn from a standard normal distribution, and the product-specific variances are drawn from a uniform distribution between one and two. Product-specific means and variances are constant over time, unless otherwise noted. Each period, the log firm appeal shifters are drawn from normal distributions with zero means and firm-specific variances. The firm-specific variances are drawn from a uniform distribution between one and two, and are constant over time. Finally, the log marginal cost shifters are drawn each period from normal distributions with zero means and product-specific variances. The product-specific variances are drawn from a uniform distribution between one and two, and are constant over time. The product appeal shifters, firm appeal shifters, and marginal cost shifters are mutually independent.

The econometrician is assumed to be able to observe the elasticities of substitution σ^U and σ^F exactly without estimation in constructing the price indices. The CES exact price indices are calculated without considering nesting of preferences among products and firms, with the exception of the unit expenditure function, which is calculated according to the consumer's exact preference structure.

do not include a burn-in period.

⁵²[Hottman et al. \(2016\)](#) found evidence of such differences and we find related evidence of differences in elasticities within and across nests. We found in section 3.4.1 this did not matter much for the properties of the CUPI. More work is needed in this area, but we do not explore this issue in our simulation analysis. Relatedly, an interesting and open question is how much the CUPI is sensitive to any biases in the estimation of the elasticities. We leave that question for future work.

Simulation Evidence

We consider five sets of simulations in this section. In each set of simulations, we vary one key parameter and run 100 Monte Carlo simulations as described in the previous section for each value of the key parameter we consider in the set of simulations. The figures display inflation as measured by the unit expenditure function and various CES price indices; the lines represent the average realization of measured inflation using each price index, while the shaded regions represent 95-percent asymptotic confidence intervals. The first three sets of simulations consider frictionless markets in which the assumptions underlying the CUPI hold exactly, so it coincides identically with the unit expenditure function in those exercises. The fourth and fifth sets of simulations introduce market imperfections that drive a wedge between the CUPI and the unit expenditure function.

Trends in Marginal Costs

Figure 3.A.16 explores the behavior of the Sato-Vartia index and CUPI in the environment of [Hottman et al. \(2016\)](#) when there is a trend in the marginal cost shifter a_{ut} . On the left-hand side of the graph, marginal costs are falling at a rate of 5 log points per period; in the middle of the graph, marginal costs have no trend; and on the right-hand side of the graph, marginal costs are rising at a rate of 5 log points per period. These trends in marginal costs drive non-zero average inflation. In this frictionless environment, the CUPI exactly replicates the unit expenditure function. The Sato-Vartia index is substantially less precise than the CUPI, as seen in its wider 95-percent simulation bands for estimated inflation. The Sato-Vartia index is noisier than the CUPI because it does not account for changes in product appeal; despite the normalization that average appeal levels are steady over time in these simulations, appeal shocks may affect the consumer's cost of living in any particular simulation. Generally speaking, if goods with large expenditure shares experience positive appeal shocks on average, the cost of living will fall, but if they experience negative appeal shocks on average, the cost of living will rise. Consistent with the logic in Section 3.9.2, however, the Sato-Vartia index does not display an average bias relative to the unit expenditure function.

Trends in Variance of Product Appeal

Figure 3.A.17 explores the behavior of the Sato-Vartia index and CUPI when there is a trend in the variance of the product appeal parameters φ_{ut} . The horizontal axis of the graph shows different growth rates for the variance of appeal; on the left-hand side of the graph, appeal is becoming more compressed over time, while on the right-hand side of the graph, appeal is becoming more dispersed over time. The unit expenditure function shows that the consumer's cost of living is falling over time when the variance of product appeal is rising, and conversely the cost of living is rising when the variance of product appeal is falling over time. This pattern is consistent with the logic in Redding and Weinstein (2020) that increasing dispersion in product appeal is valuable to consumers when products are substitutes, because it provides greater opportunities for substitution to preferred varieties. In contrast to the results in Figure 3.A.16, the Sato-Vartia index does exhibit an average bias in the presence of time trends in the variance of product appeal, which is especially evident in the right-hand portion of the figure where the variance is growing over time. This figure helps illustrate the potential benefits of using the CUPI.

As noted above, however, rising dispersion in product appeal seems empirically unable to account for the rapid deflation implied by the CUPI. Figure 3.A.5 shows that several product groups in the NPD data exhibit nonmonotonic patterns in measured appeal dispersion, even where there is a large gap between the CUPI and Feenstra index. We therefore conclude that rising dispersion in product appeal is unlikely to account for the empirical gap between the CUPI and the other exact CES price indices.

Product Upgrading and Downgrading via Turnover

Figure 3.A.18 displays results from simulations featuring product entry and exit. For simplicity, we assume that products are present in the market place for a deterministic number of periods (set to five in these simulations) after which they exit. Equal numbers of products enter and exit the market in every period.

The key feature of the simulations is that the average appeal parameter φ_{ut} for entering

products can differ from the average for continuing products.⁵³ The horizontal axis of the graph shows different trends in the average appeal of entering products. On the left-hand side of the graph, entering products are less appealing on average than existing products, while on the right-hand side of the graph, entering products are more appealing.

Figure 3.A.18 shows inflation as measured by the Sato-Vartia index, the CUPI, and the Feenstra index, which is equal to the Sato-Vartia index in the absence of product entry and exit. The CUPI again tracks the true unit expenditure function exactly, showing inflation from product downgrading and deflation from product upgrading. The Sato-Vartia index captures these effects directionally, because product turnover affects the prices of continuing products via competition. Because it considers only continuing products, however, the Sato-Vartia index quantitatively understates product turnover's effects on the cost of living. The Feenstra index augments the Sato-Vartia index with an adjustment term that captures the effects of product turnover directly. Figure 3.A.18 shows that it is unbiased on average relative to the true unit expenditure function, despite the presence of relative appeal shocks in the simulations. Echoing the results of Figure 3.A.16, the Sato-Vartia index and the Feenstra index are noisier than the CUPI because they do not account for the effect of product appeal shocks. This figure helps make the case for using an index such as the Feenstra or CUPI to incorporate product turnover that yields quality change.

Segmented Markets

Figure 3.A.19 displays results from a set of simulations in which the market is segmented into five distinct submarkets. Consumers have nested CES preferences over the products consumed in each submarket and firms compete within each submarket as described in Section 3.9.3. Consumers have Cobb-Douglas preferences over their consumption across the various submarkets. One of the markets is “large,” and has a weight of 0.8 in the consumer's aggregate utility function, while the other four markets are “small,” and have weights of 0.05 each. Product entry and exit within each market otherwise proceeds as in the previous set of simulations.

⁵³Recall that the normalization on product appeal in Redding and Weinstein (2020) applies only to continuing products, so product upgrading or downgrading does not violate the normalization.

The simulations present price indices measured assuming that the econometrician is unaware of the market segmentation and measures prices assuming a unified marketplace. The assumptions are meant to mimic the pattern documented in Figure 3.A.6, which shows that although most sales are concentrated among products sold in nearly all metro areas nationally, on a UPC basis, most products are sold in relatively few areas.

As in Figure 3.A.18, the horizontal axis of Figure 3.A.19 shows different trends in the average appeal parameter φ_{it} of entering products. Only the small markets feature a trend in the average appeal of entering products; there is no trend in the large market. Figure 3.A.18 displays inflation as measured by five price indices in addition to the unit expenditure function: the Sato-Vartia; the Feenstra; the CUPI with no common goods rule, which we have called the “theoretical CUPI”; the CUPI implemented with a 40th-percentile common goods rule; and the CUPI implemented with an 80th-percentile common goods rule.

Figure 3.A.19 conveys a few key messages. First, the theoretical CUPI is significantly biased in the presence of product upgrading or downgrading in the small markets. The intuition for this bias is that the P^* and S^* ratio terms in the CUPI are unweighted geometric means. The theoretical CUPI therefore assigns the price movements in the small markets, driven by product turnover, equal importance to the price movements in the large market. Although that equal weighting scheme would be theoretically justified in a unified marketplace under CES preferences with appeal shocks, it implicitly overweights the small markets in the segmented market environment. The second key message is that the Sato-Vartia and Feenstra indices fare better in these simulations than the theoretical CUPI because all of their components are expenditure-share weighted. The third key message is that a common goods rule (CGR) can help reduce the bias in the theoretical CUPI by reallocating products from the unweighted geometric mean terms to the lambda ratio term, which is weighted.

Figure 3.A.19 thus provides a theoretical justification for the use of a CGR in Redding and Weinstein (2020) and our own empirical work. We interpret this segmented markets case as broadly capturing the intuition for a CGR given that goods may first enter local markets. While this exercise helps justify a CGR, it highlights that choosing the appropriate CGR will depend on the pace of product upgrading and degree of market segmentation.

In addition, in practice entering goods can transition to becoming national goods, and that process will influence the nature of the CGR. Put differently, although this exercise provides theoretical motivation for a CGR, it does not provide precise guidance as to the nature of the appropriate CGR.

Partial Stock-outs (Rationing) Prior to Exit

Figure 3.A.20 examines the behavior of the CES exact price indices when there are partial product stock-outs in the period prior to exit. The simulations feature a stylized version of stock-outs, or a “clearance rack,” in which product sales are rationed in the period before they exit the marketplace. Product entry and exit within each market otherwise proceeds as in the previous two sets of simulations.

The horizontal axis of the figure shows various shares of rationing prior to exit. On the left-hand side of the figure, consumers are only able to purchase 10 percent of their desired (unconstrained) product demands; on the right-hand side of the figure, there is no rationing. We assume that firms do not adjust stocked-out products’ prices to clear the market, but instead price all products as they would in the flexible price equilibrium. We assume that consumers optimally reallocate their demands toward the unconstrained products in response to the rationing.⁵⁴

The unit expenditure function in Figure 3.A.20 shows an approximately constant cost of living in the presence of stock-outs. Although the simulations feature product turnover, they do not feature any trend in average appeal of entering products. As the figure shows, though, stock-outs introduce a substantial bias to the CUPI and the Feenstra index. The intuition for the bias is subtle. Rationing lowers expenditure shares on goods just prior to their exit from the marketplace, with the expenditure reallocated to unconstrained goods. Rationing therefore raises the dispersion of expenditure shares on continuing goods relative to the un-rationed case, leading to a negative $\log S^*$ ratio.⁵⁵ Likewise, the Feenstra adjust-

⁵⁴The assumption that consumers have homothetic CES preferences makes it straightforward to calculate their re-optimized demands in the presence of rationing; consumers reallocate their expenditure to each of the non-rationed goods in proportion to their unconstrained demands had there been no rationing. The unit expenditure function under rationing can then be computed as the ratio of indirect utilities provided by a unit of expenditure between periods.

⁵⁵The entry of the unconstrained goods does not affect this calculation, because the expenditure shares in

ment to the Sato-Vartia index is negative because new goods enter the market un-rationed, allowing consumers to buy whatever quantities they please; prior to exit, quantities are constrained below consumers' desired levels. The expenditure share on exiting products is therefore lower than the expenditure share on entering products, producing a negative adjustment to both the Feenstra index and the CUPI.

The key messages from Figure 3.A.20 are similar to those from Figure 3.A.19. The theoretical CUPI is significantly biased in the presence of this market friction, while the Sato-Vartia index is approximately unbiased. Imposing a CGR helps move the CUPI closer to the true unit expenditure function. Again, though, the simulations do not provide guidance on the empirically appropriate CGR. Estimates of the extent and nature of rationing are needed to yield guidance for the appropriate CGR.

the CUPI's consumer valuation adjustment term are calculated over continuing goods only.

3.9.4 Appendix Tables and Figures

Table 3.A.1: Impact of Alternative Imputation for Missing Price of Entrants, NPD Data

	Memory Cards	Coffeemakers	Headphones	Boys' Jeans	Occupational Footwear
Hedonic Tornqvist (TV,EP2)	-20.12	-9.56	-14.13	-9.16	-3.79
Hedonic Tornqvist (TV, EP2H)	-20.06	-9.57	-13.74	-8.92	-3.78

Notes: EP2 uses the (Erickson and Pakes, 2011) method imputing the missing price relative for entrants assuming the lagged residual is zero. EP2H imputes the missing price relative using the current period residual. Reported are annual averages of chained price indices for Q4. Data come from the NPD Group.

Table 3.A.2: Comparison of GEKS-Lite and Rolling Year GEKS, Annual Chained Price Indices, NPD Data

	Memory Cards	Coffeemakers	Headphones	Boys' Jeans	Occupational Footwear
Tornqvist(C)	-17.34	-8.48	-10.93	-7.35	-3.29
Tornqvist(GL)	-15.89	-6.42	-10.91	-5.40	-2.28
Tornqvist(RYGEKS)	-15.89	-6.47	-8.72	-4.66	-1.94

Notes: C is chained. GL is Geks-Lite. RYGEKS is rolling year GEKS. Reported are annual average price indices for Q4. Data come from the NPD Group.

Table 3.A.3: Summary Statistics for Alternative Price Chained Indices, GEKS-Lite, NPD Data

	Memory Cards	Coffeemakers	Headphones	Boys' Jeans	Occupational Footwear
Mean(Tornqvist)	-15.41	-6.64	-11.55	-5.55	-2.31
SD (Tornqvist)	10.85	4.13	4.20	3.40	3.05
Mean(Tornqvist(TV,QW))	-20.60	-10.06	-14.51	-7.94	-3.76
SD (Tornqvist(TV,QW))	8.57	4.53	5.39	3.23	3.28
Mean(Tornqvist(TV,EW))	-16.89	-8.21	-11.07	-6.72	-3.14
SD(Tornqvist(TV,EW))	9.45	4.60	5.05	3.20	3.01
Mean(Sato-Vartia)	-14.32	-6.36	-11.34	-4.13	-2.08
SD(Sato-Vartia)	11.03	4.25	4.15	2.85	3.06
Mean(Feenstra)	-16.46	-9.43	-13.06	-5.51	-3.76
SD(Feenstra)	9.69	4.09	5.35	2.88	3.46
Difference(Tornqvist,TV(EW))	1.48	1.57	-0.48	1.17	0.83
Difference(Tornqvist,TV(QW))	5.19	3.42	2.96	2.39	1.45
Difference(Sato-Vartia,Feenstra)	2.14	3.07	1.72	1.38	1.68
Corr(Tornqvist,Tornqvist(TV,QW))	0.99	0.91	0.96	0.97	0.99
Corr(Tornqvist,Tornqvist(TV,EW))	1.00	0.84	0.98	0.98	0.98
Corr(Tornqvist(TV,QW),Tornqvist(TV,EW))	1.00	0.97	1.00	1.00	1.00
Corr(Feenstra,Tornqvist)	0.97	0.98	0.99	0.98	0.93
Corr(Feenstra,Tornqvist(TV,QW))	0.96	0.83	0.98	0.96	0.97
Corr(Feenstra,Tornqvist(TV,EW))	0.97	0.71	0.95	0.97	0.98
Corr(Sato-Vartia,Tornqvist)	0.99	1.00	0.98	1.00	1.00
Corr(Sato-Vartia,Tornqvist(TV,QW))	0.97	0.87	0.99	0.94	0.99
Corr(Sato-Vartia,Tornqvist(TV,EW))	0.98	0.80	0.94	0.96	0.98
Corr(Sato-Vartia, Feenstra)	0.99	0.98	1.00	0.97	0.92

Notes: GEKS-lite is the average of the geometric mean of the chained values and the YoY price indices for q4 for each year. QW uses quantity=weights in the TV hedonic estimation. EW uses expenditure weights. Data come from the NPD Group.

Table 3.A.4: Hedonic Models: Goodness of Fit

R^2 for:	Log Price Level		Log Price Relative			
Model:	Log-Level	Log-Level	EP1		EP2	
Coffee Makers	0.62	0.05	0.20	0.20	0.21	0.23
Headphones	0.89	0.24	0.47	0.47	0.51	0.49
Memory Cards	0.71	0.05	0.10	0.09	0.15	0.13
Work/Occ Footwear	0.73	0.1	0.33	0.40	0.34	0.41
Boy's Jeans	0.72	0.22	0.36	0.46	0.42	0.50
Weights:	QW	QW	EW	QW	EW	QW

Average quarterly R^2 for hedonic regression models. For cases where the outcome variable (price level or price relative) does not match the LHS variable from the hedonic model, we report the R^2 from a regression of transformed predicted values from the hedonic model on actual values. For example, the price-relative R^2 for the log-level model is the R^2 from a regression of price relatives constructed from a log-level hedonic model on actual price relatives. Weights used in regressions are consistent in hedonic estimation and construction of R^2 measures. For the log-level model, weights are the quantity shares (QW) in the current period. For the log-difference and time-varying unobservables model, weights are the average quantity shares (QW) or the average expenditure shares (EW) in the current and lagged periods. The time-varying unobservable model includes lagged residuals from a log-level hedonic regression.

Table 3.A.5: Nested Estimated Elasticities of Substitution: NPD Data

Product	Groups	Elasticity of Substitution			
		Within		Across	
Headphones	Manual	8.609	(0.544)	7.704	(0.491)
	Hedonic	9.537	(0.969)	8.958	(0.423)
Memory Cards	Manual	6.31	(0.675)	4.534	(0.298)
	Hedonic	6.621	(0.657)	5.25	(0.586)
Coffeemakers	Manual	5.495	(0.791)	3.42	(0.63)
	Hedonic	5.345	(0.99)	5.306	(0.374)
Occupational Footwear	Manual	5.545	(0.509)	3.057	(0.493)
	Hedonic	6.199	(0.548)	4.135	(0.769)
Boys' Jeans	Manual	7.439	(1.5)	3.234	(0.734)
	Hedonic	8.156	(1.82)	3.418	(0.657)

Notes: Estimated elasticities of substitution for nested CES models. Standard errors in parentheses. Data come from NPD Group. Within-nest elasticities are estimated using the methodology of [Feenstra \(1994\)](#) and [Redding and Weinstein \(2020\)](#). Across-nest elasticities are estimated using the nested CES estimation procedure of [Hottman et al. \(2016\)](#) modified to be robust to product entry and exit.

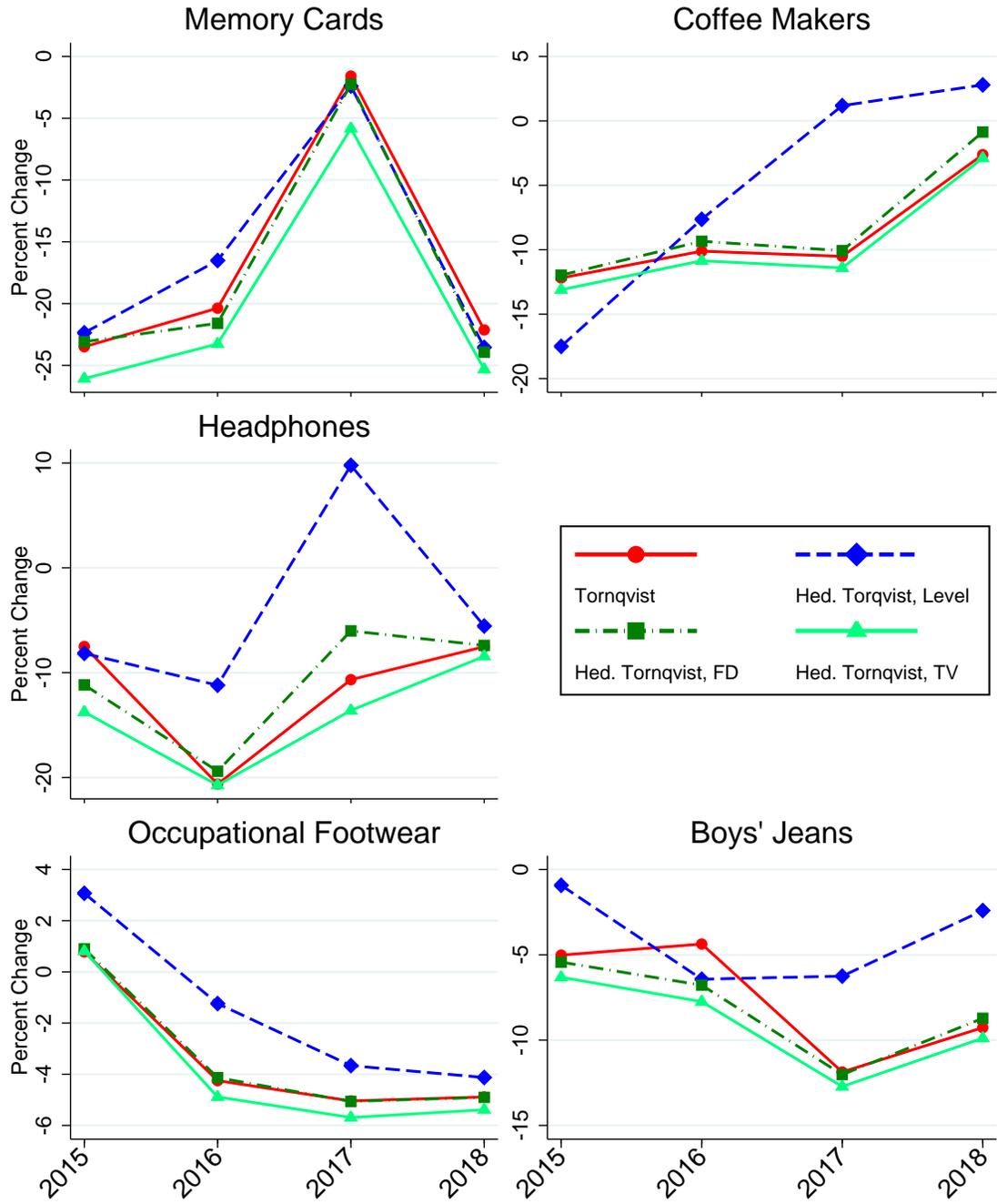
Table 3.A.6: Food Product Groups: Nielsen Retail Scanner Data

Product Group	Product Group Code	Product Group	Product Group Code
Baby Food	501	Juice, Drinks - Canned, Bottled	507
Baked Goods-Frozen	2001	Juices, Drinks-Frozen	2006
Baking Mixes	1001	Milk	2506
Baking Supplies	1002	Nuts	1011
Bread And Baked Goods	1501	Packaged Meats-Deli	3002
Breakfast Food	1004	Packaged Milk And Modifiers	1012
Breakfast Foods-Frozen	2002	Pasta	1013
Butter And Margarine	2501	Pickles, Olives, And Relish	1014
Candy	503	Pizza/Snacks/Hors D'oeuvres-Frzn	2007
Carbonated Beverages	1503	Prepared Food-Dry Mixes	511
Cereal	1005	Prepared Food-Ready-To-Serve	510
Cheese	2502	Prepared Foods-Frozen	2008
Coffee	1006	Pudding, Desserts-Dairy	2507
Condiments, Gravies, And Sauces	1007	Salad Dressings, Mayo, Toppings	1015
Cookies	1505	Seafood - Canned	512
Cot Cheese, Sour Cream, Toppings	2503	Shortening, Oil	1016
Crackers	1506	Snacks	1507
Desserts, Gelatins, Syrup	1008	Snacks, Spreads, Dips-Dairy	2508
Desserts/Fruits/Toppings-Frozen	2003	Soft Drinks-Non-Carbonated	1508
Dough Products	2504	Soup	513
Dressings/Salads/Prep Foods-Deli	3001	Spices, Seasoning, Extracts	1017
Eggs	2505	Sugar, Sweeteners	1018
Flour	1009	Table Syrups, Molasses	1019
Fresh Meat	3501	Tea	1020
Fresh Produce	4001	Unprep Meat/Poultry/Seafood-Frzn	2009
Fruit - Canned	504	Vegetables - Canned	514
Fruit - Dried	1010	Vegetables And Grains - Dried	1021
Gum	505	Vegetables-Frozen	2010
Ice Cream, Novelties	2005	Yogurt	2510
Jams, Jellies, Spreads	506		

Table 3.A.7: Nonfood Product Groups: Nielsen Retail Scanner Data

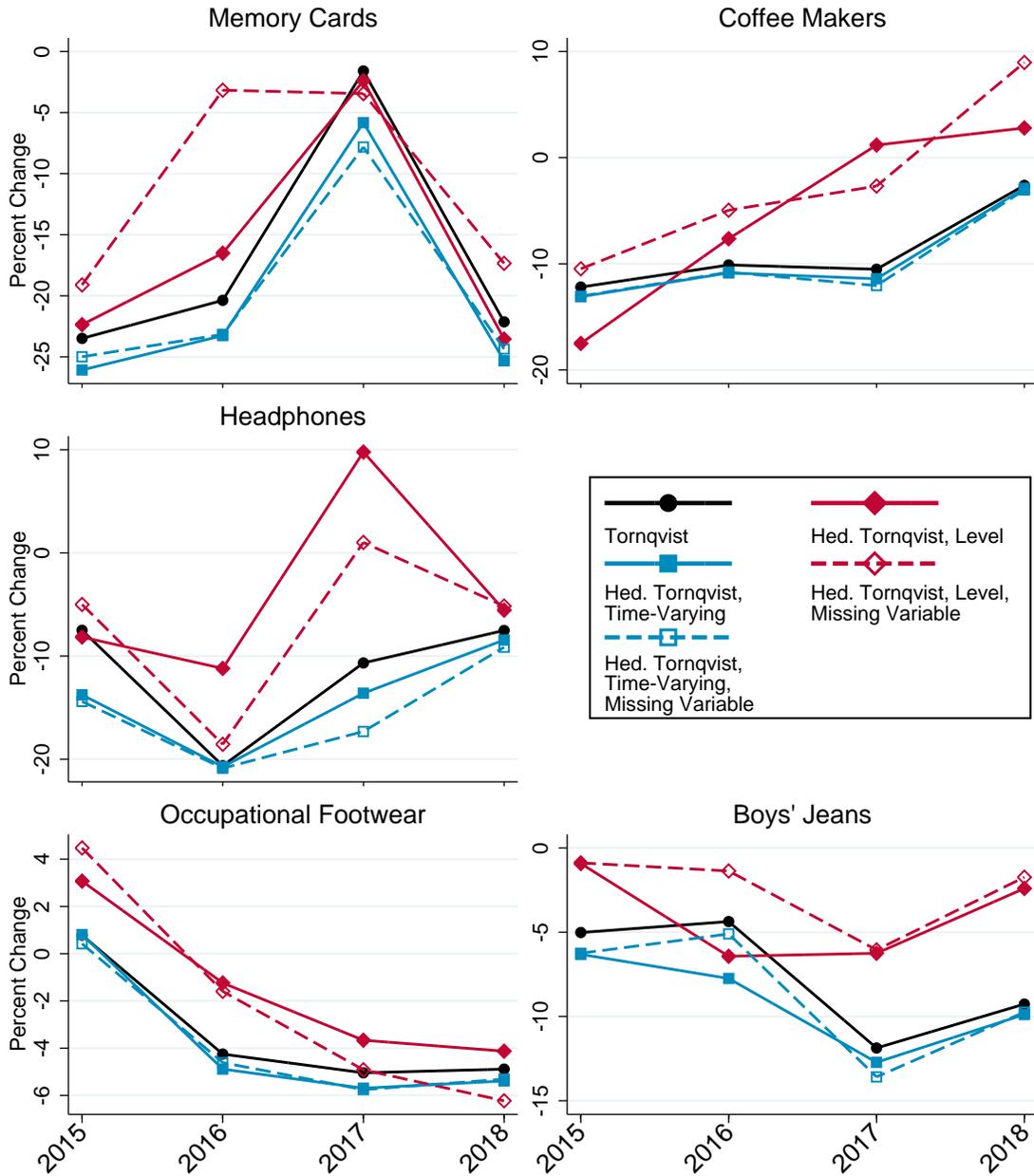
Product Group	Product Group Code	Product Group	Product Group Code
Automotive	5501	Housewares, Appliances	5513
Baby Needs	6001	Ice	2004
Batteries And Flashlights	5502	Insecticds/Pesticds/Rodenticds	5514
Beer	5001	Kitchen Gadgets	5515
Books And Magazines	5503	Laundry Supplies	4506
Canning, Freezing Supplies	5504	Light Bulbs, Electric Goods	5516
Charcoal, Logs, Accessories	5505	Liquor	5002
Cosmetics	6002	Medications/Remedies/Health Aids	6012
Cough And Cold Remedies	6003	Men's Toiletries	6013
Deodorant	6004	Oral Hygiene	6014
Detergents	4501	Paper Products	4507
Diet Aids	6005	Personal Soap And Bath Additives	4508
Disposable Diapers	4502	Pet Care	4509
Electronics, Records, Tapes	5507	Pet Food	508
Ethnic Haba	6006	Photographic Supplies	5517
Feminine Hygiene	6007	Sanitary Protection	6015
First Aid	6008	Sewing Notions	5519
Floral, Gardening	5508	Shaving Needs	6016
Fragrances - Women	6009	Shoe Care	5520
Fresheners And Deodorizers	4503	Skin Care Preparations	6017
Glassware, Tableware	5509	Stationery, School Supplies	5522
Grooming Aids	6010	Tobacco & Accessories	4510
Hair Care	6011	Vitamins	6018
Hardware, Tools	5511	Wine	5003
Household Cleaners	4504	Wrapping Materials And Bags	4511
Household Supplies	4505		

Figure 3.A.1: Alternative Hedonic Estimation Strategies NPD Data



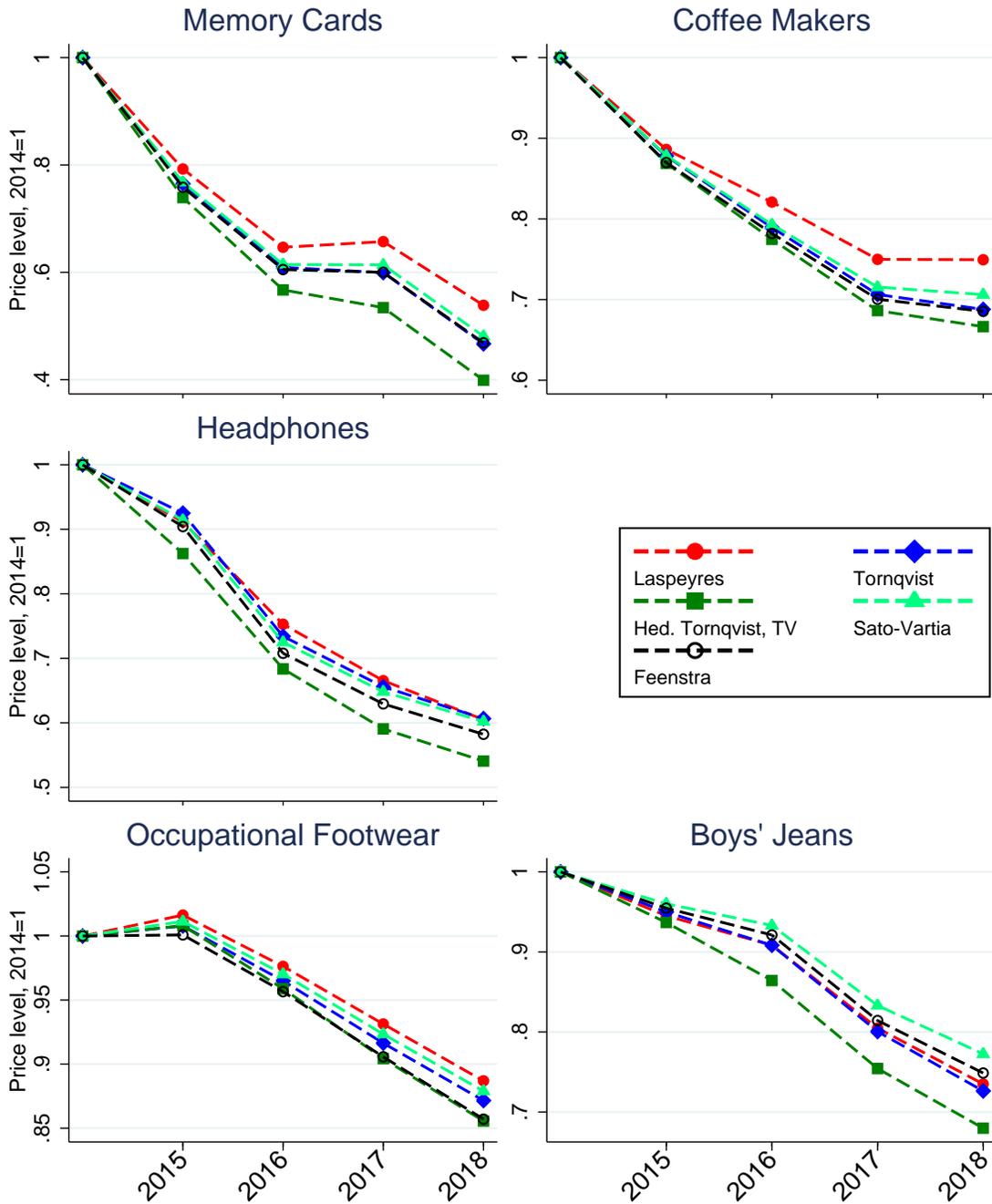
Notes: Values are percent change on a q4-to-q4 basis, aggregated from chained quarterly indices. Data comes from the NPD Group.

Figure 3.A.2: Test of Time-Varying Unobservable Hedonic Specification, First-Differences and Levels Estimation



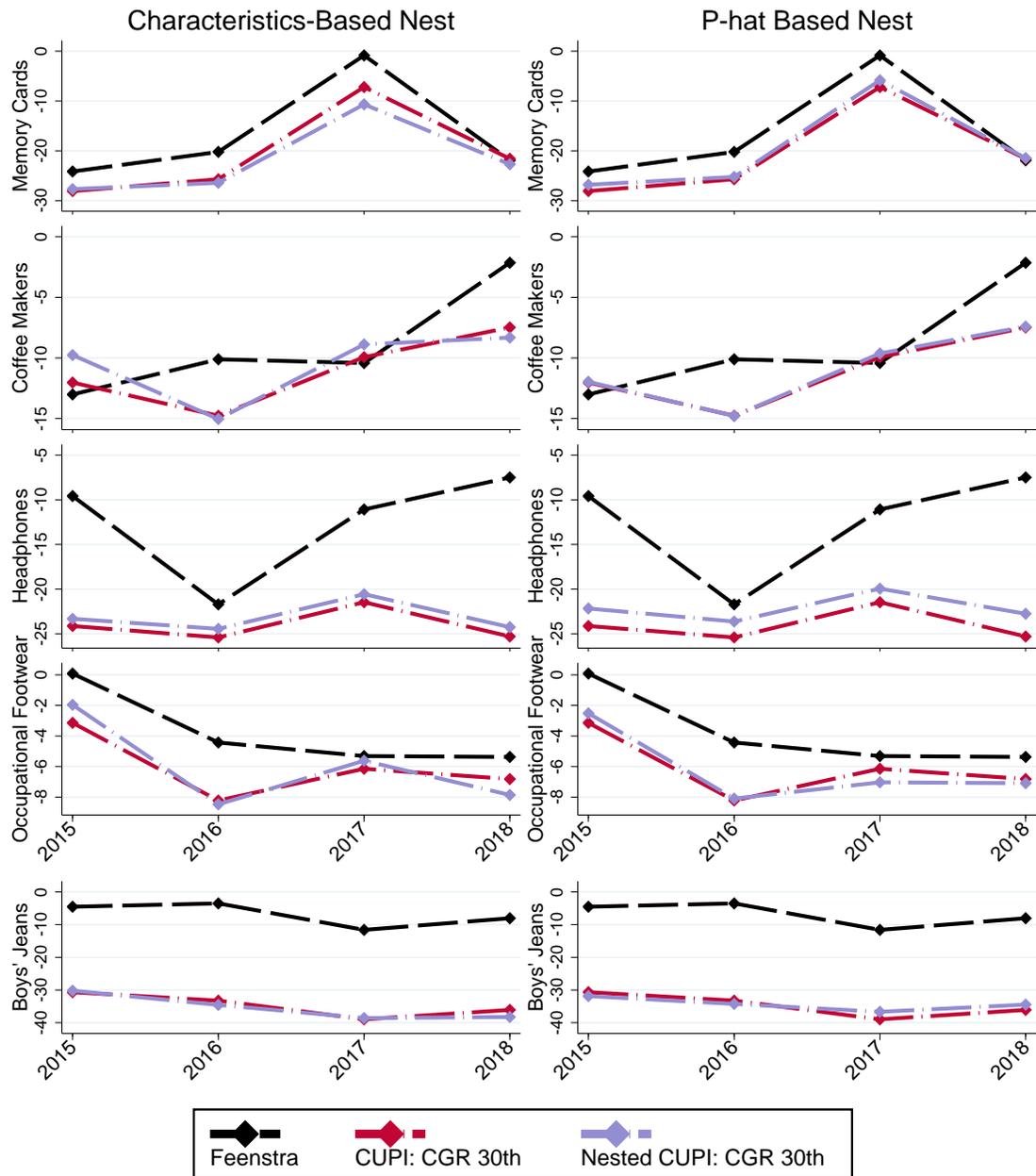
Notes: Values are percent change on a q4-to-q4 basis, aggregated from chained quarterly indices. Data comes from the NPD Group.

Figure 3.A.3: Main Price Index Specifications: Cumulative Price Level Changes, No CUPI



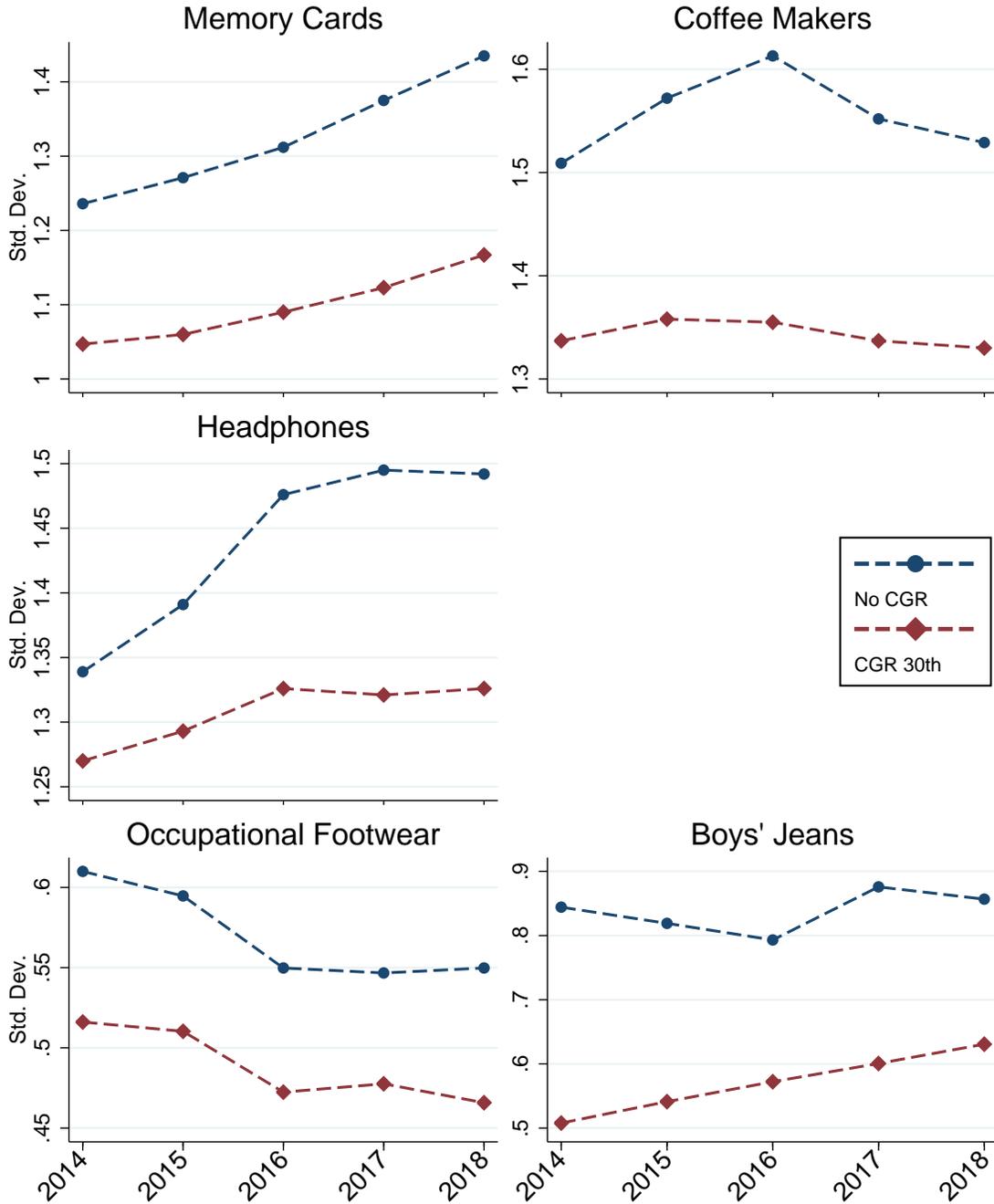
Notes: Values are cumulative changes relative to the 2014 price level, with 2014 price level set to 1. The hedonic time-varying unobservables model is estimated over log price differences using WLS and with weights that are average quantity-shares in adjacent periods. Data comes from the NPD Group.

Figure 3.A.4: Nested CUPI: Characteristics- and P-Hat- Based Nests Percent Changes



Notes: Values are percent change on a q4-to-q4 basis, aggregated from chained quarterly indices. Laspeyres is the geometric Laspeyres. For the characteristics-based nests, we assign items to groups based on shared observable characteristics. The p-hat based nests are based on the decile of predicted prices from unweighted hedonic log-level models. We estimate period-by-period hedonic models and assign items their most common decile over all periods. Data comes from the NPD Group.

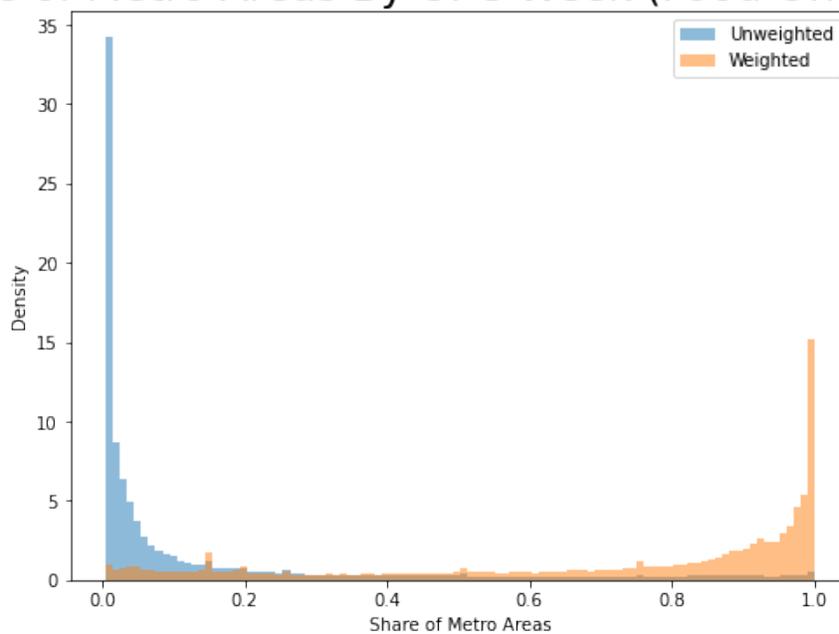
Figure 3.A.5: Dispersion of Relative Product Appeal



Notes: Values are annual averages of quarterly dispersion (standard deviation) in normalized log relative product appeal for common goods. Reported are the annual averages without imposing a common goods rule and also those with a 30th percentile common goods rule. Data comes from the NPD Group.

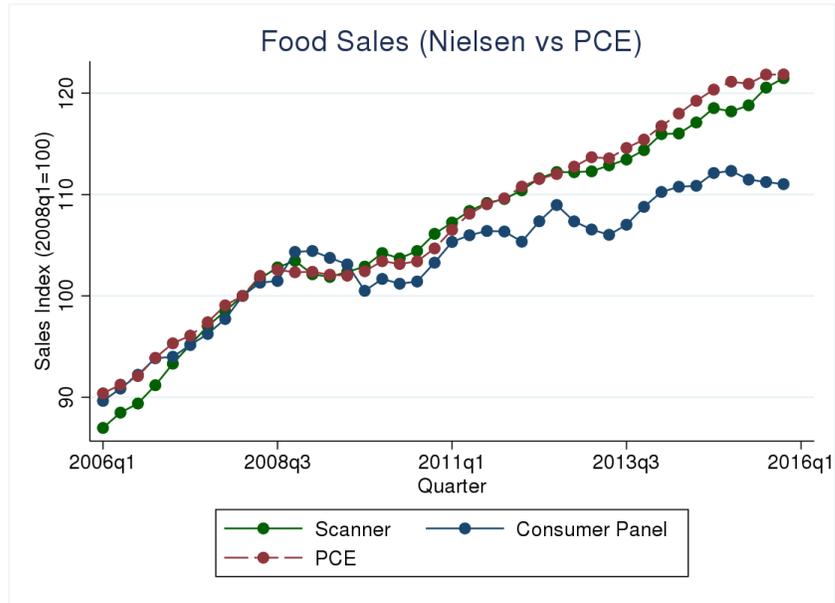
Figure 3.A.6: Sales-weighted and Unweighted Distributions of Market Penetration of Items in Nielsen Data, Food

Share of Metro Areas By UPC-Week (Food Only, 2007)

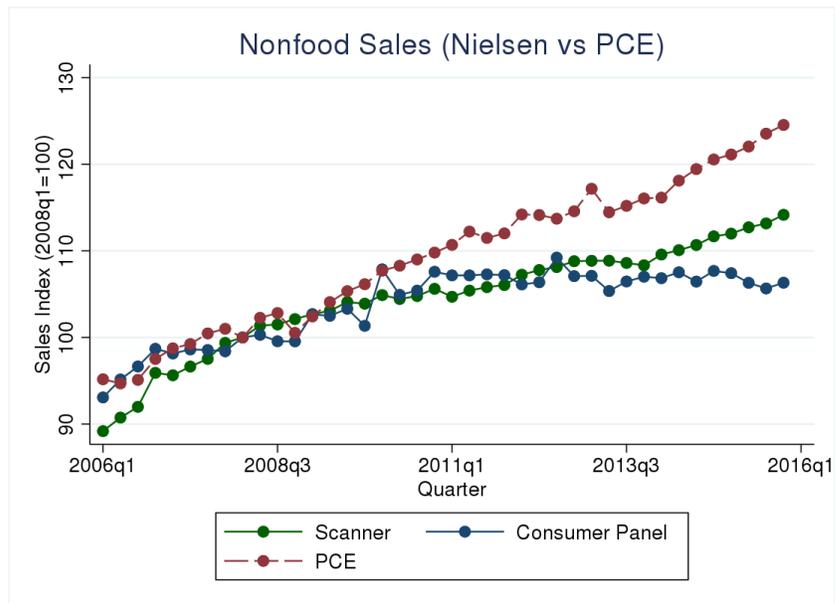


Notes: All UPC items at a weekly frequency are used from 2006-2015. Unweighted shows the market penetration at the metro area of the unweighted pooled distribution. Sales-weighted shows the equivalent using sales weights. Figure uses Nielsen Retail Scanner data for food product groups.

Figure 3.A.7: PCE vs Nielsen Sales for Scanner and Consumer Panel, Food and Nonfood
(a) Food

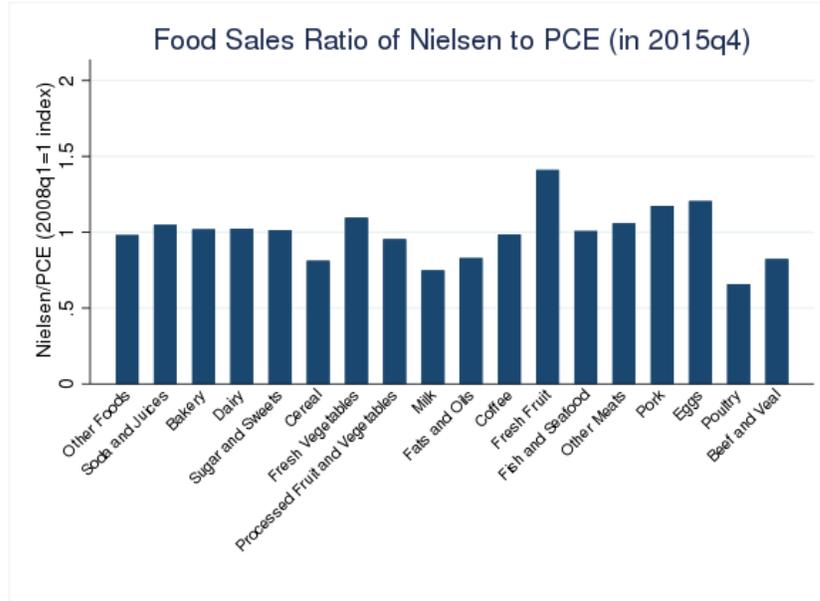


(b) Nonfood

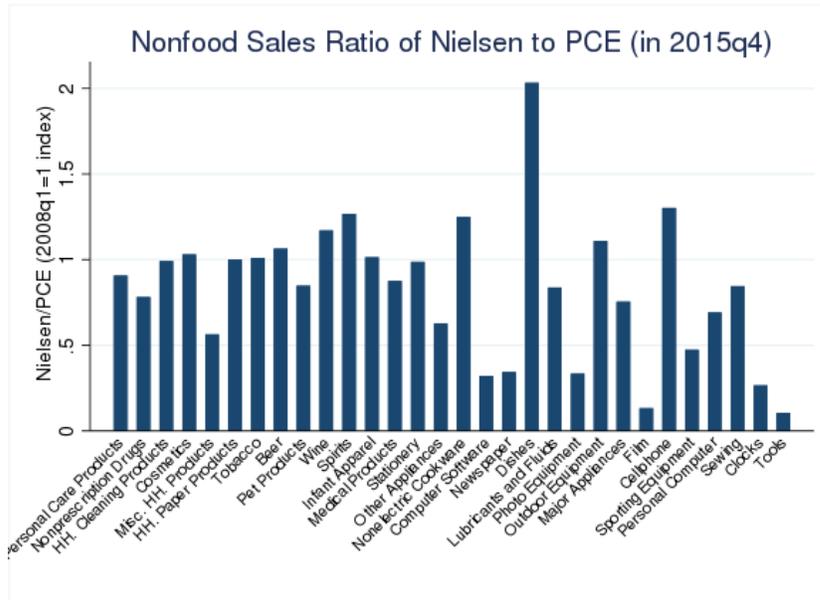


Notes: Figures uses Nielsen Scanner and Consumer Panel data for Food and Nonfood (aggregated) product groups. PCE is personal consumption expenditures (nominal) from BEA. All series indexed to 1 in 2008:q1.

Figure 3.A.8: PCE vs Nielsen Sales for Scanner, By Category within Food and Nonfood
(a) Food

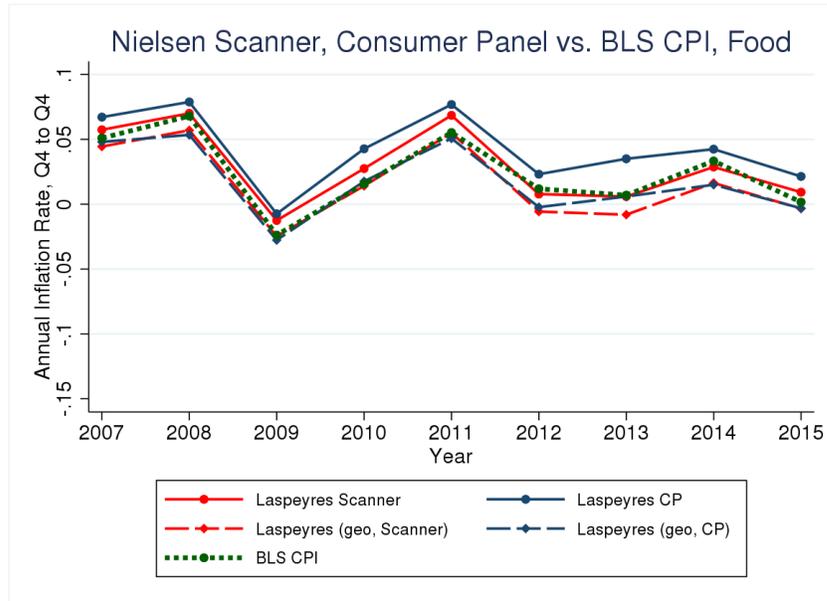


(b) Nonfood

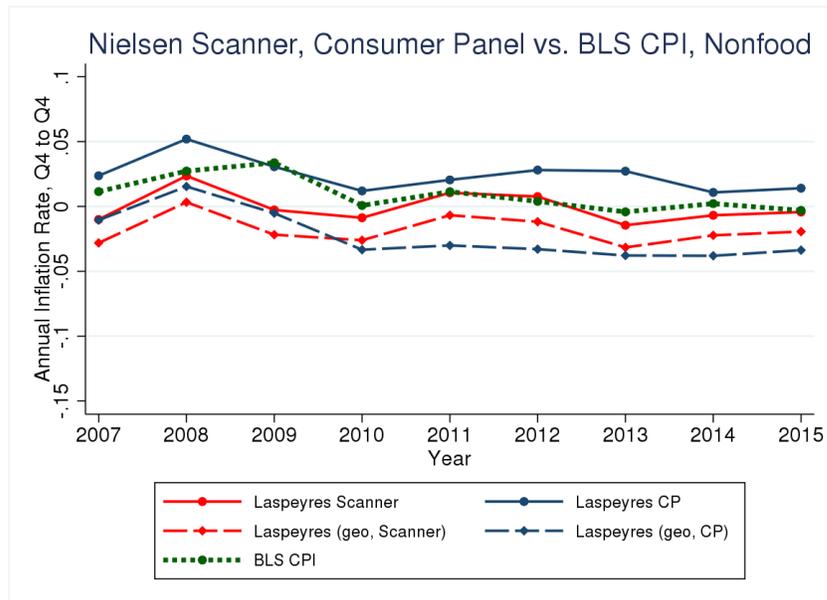


Notes: Figures uses Nielsen Scanner and Consumer Panel data for Food and Nonfood (aggregated) product groups. PCE is personal consumption expenditures (nominal) from BEA. All series indexed to 1 in 2008:q1.

Figure 3.A.9: BLS CPI vs Nielsen Laspeyres, Food and Nonfood
(a) Food

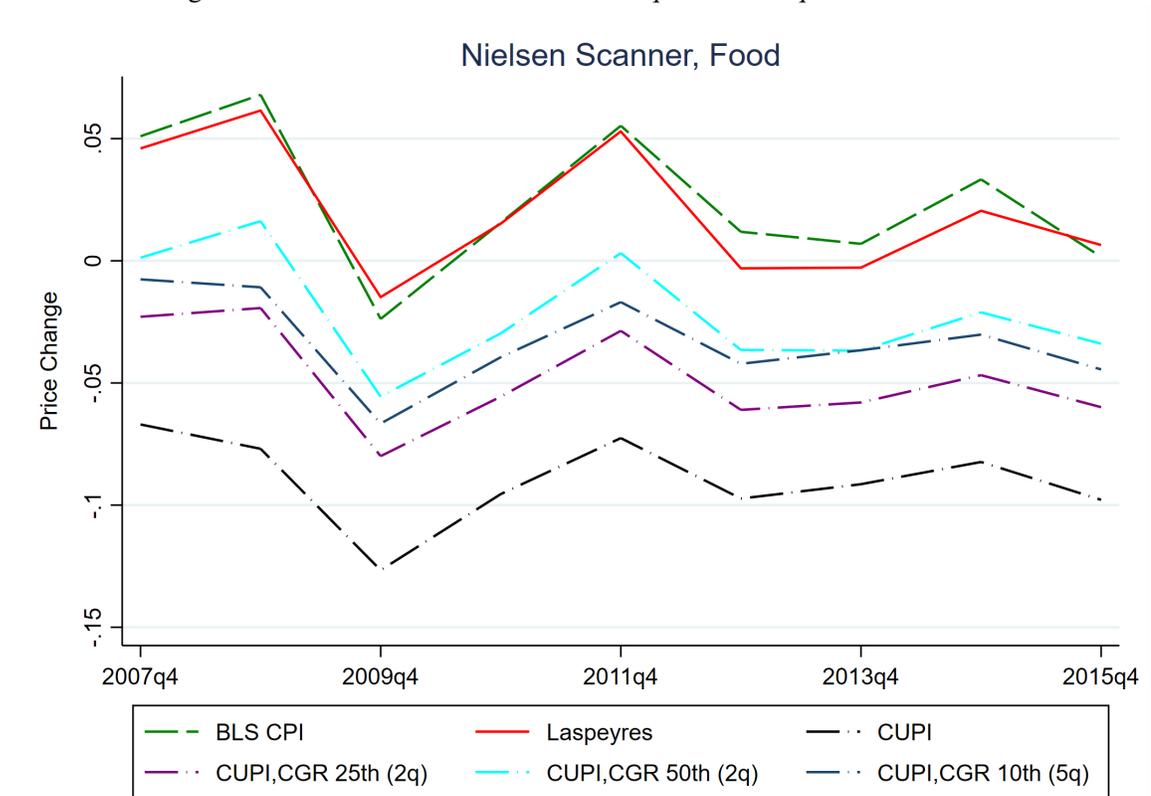


(b) Nonfood



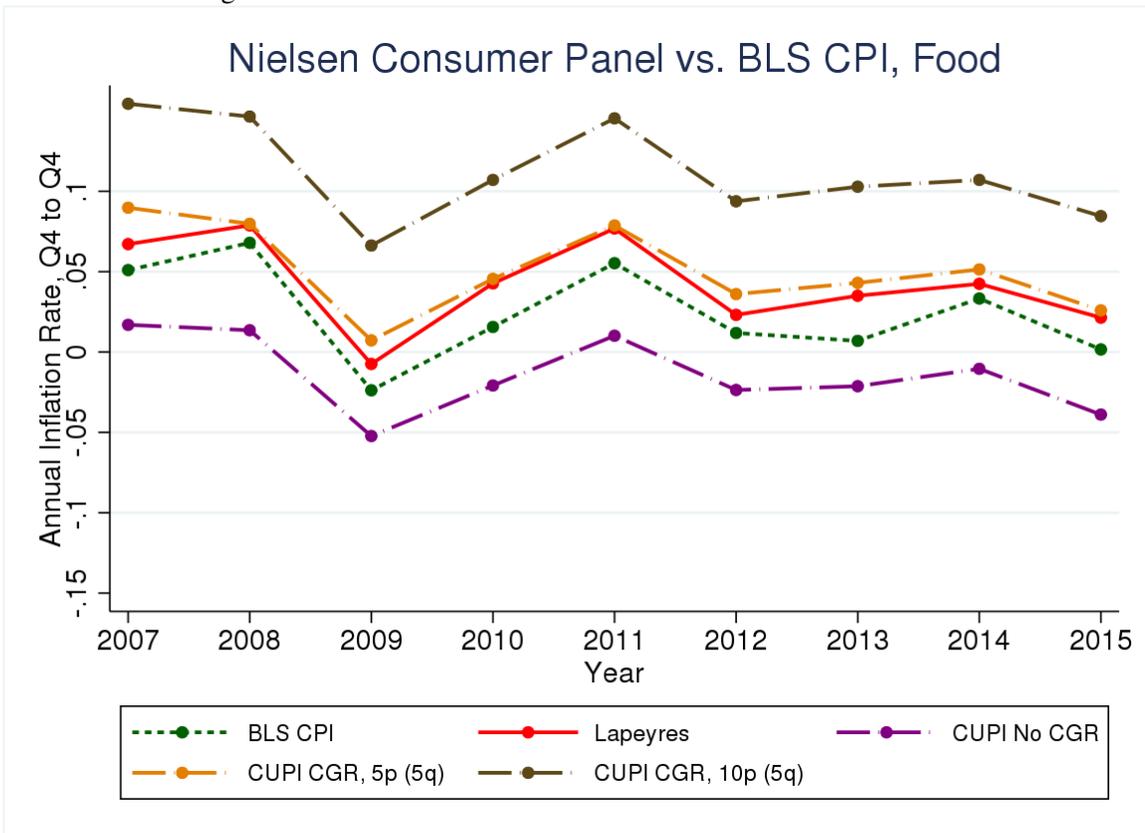
Notes: The figures show annual inflation from the fourth quarter of the previous year to the fourth quarter of the labeled year calculated from chained quarterly price indices. The panels use Nielsen Scanner and Consumer Panel data for Food and Nonfood (aggregated) product groups. The BLS CPI was computed by BLS staff.

Figure 3.A.10: Common Goods Rules–2-quarter vs 5-quarter Horizons



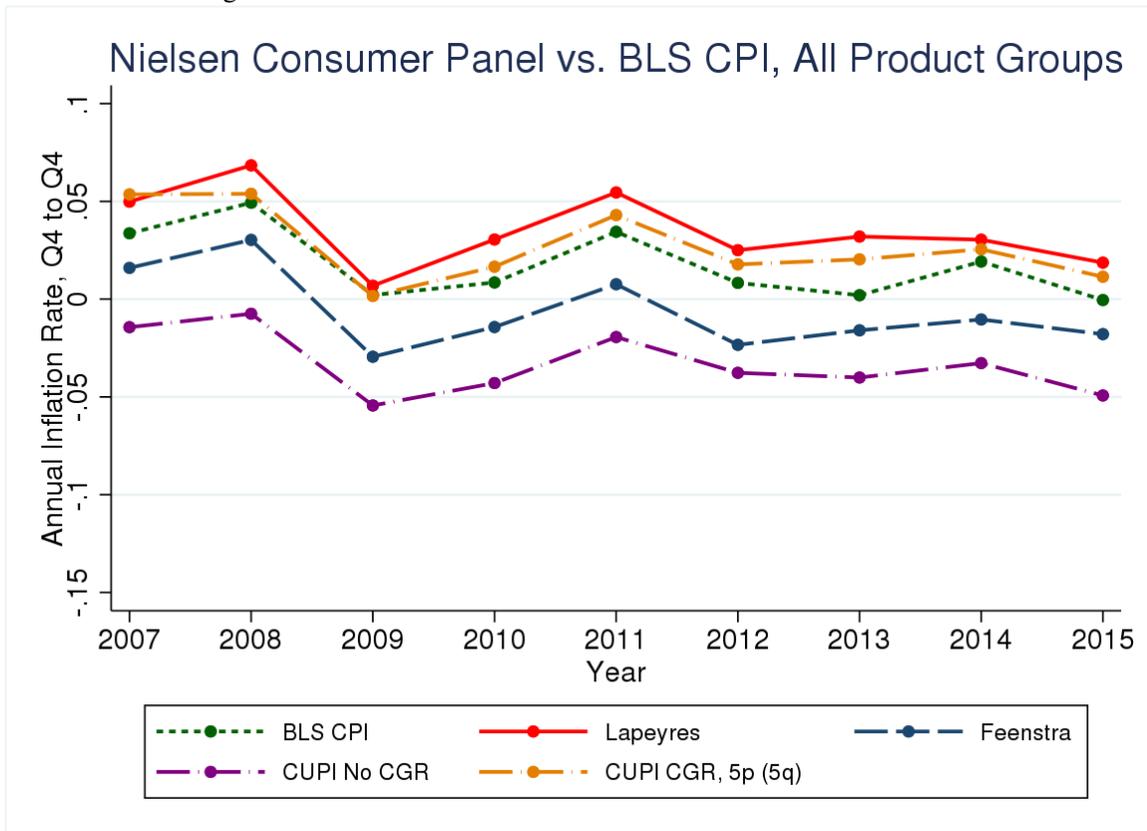
Notes: Figure uses Nielsen Scanner data for food. The 2q CUPI computes CGR percentile thresholds using sales pooled over a two quarter horizon (t and $t - 1$). The 5q CUPI computes CGR percentile thresholds using sales pooled over a 5 quarter horizon (current and prior 4 quarters). Laspeyres is arithmetic.

Figure 3.A.11: Common Goods Rules – Nielsen Consumer Panel



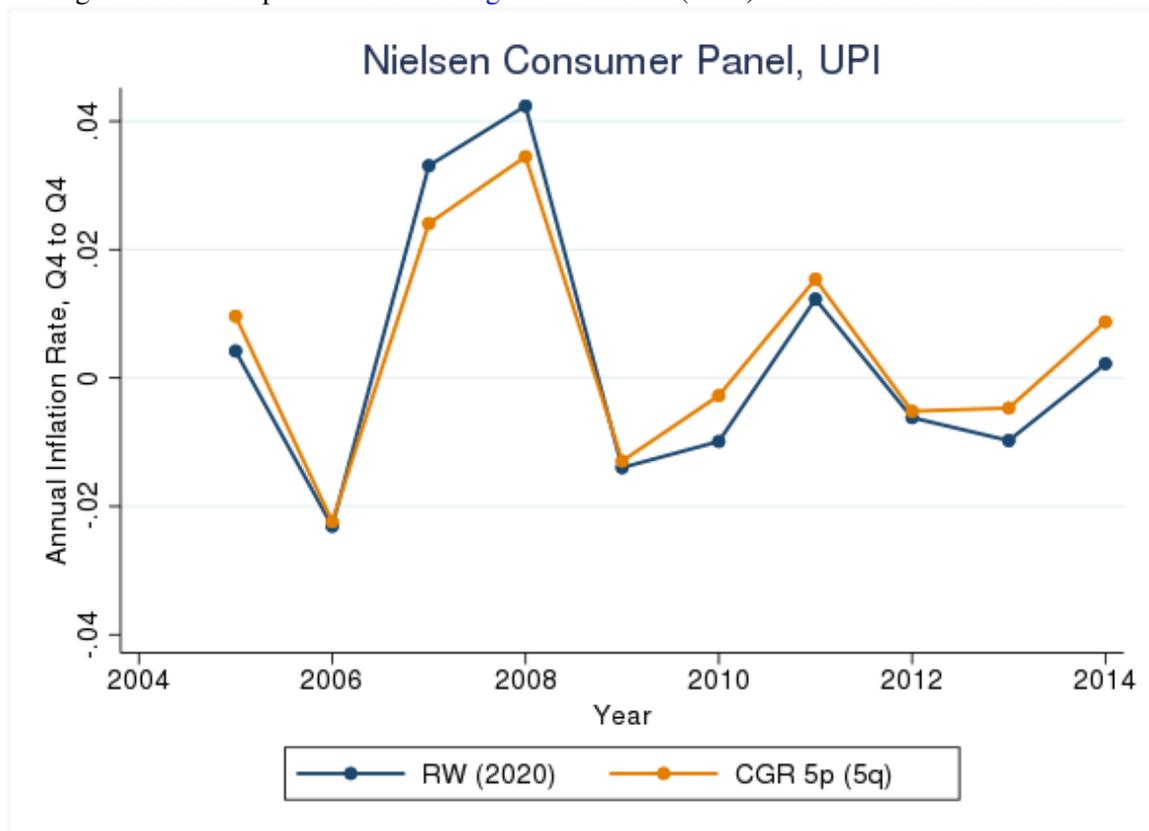
Notes: Figure uses Nielsen Consumer Panel data for food. The 5q CUPI computes CGR percentile thresholds using sales pooled over a five quarter horizon (t and $t - 1$). (current and prior 4 quarters). Laspeyres is arithmetic.

Figure 3.A.12: Common Goods Rules – Nielsen Consumer Panel



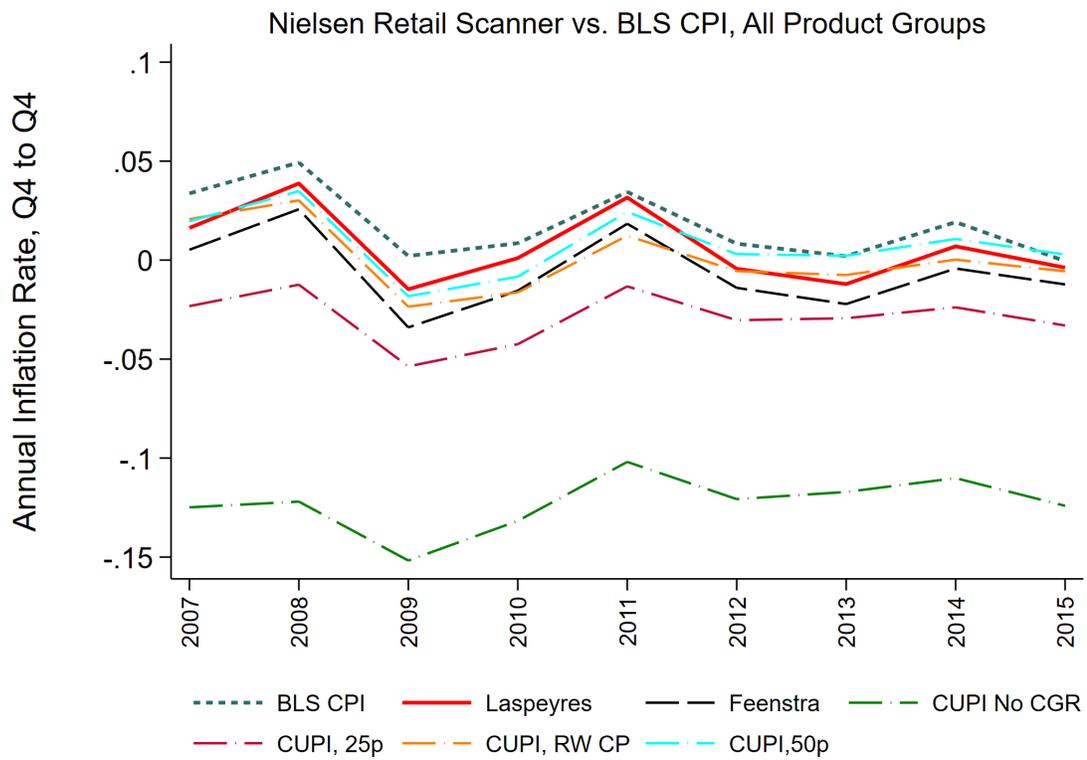
Notes: Figure uses Nielsen Consumer Panel data for food and nonfood product groups. The series “CUPI CGR RW” uses a 5th-percentile sales cutoff for the common goods rule. Percentile computed from sales pooled over 5 quarter horizon (current and prior 4 quarters). Laspeyres is arithmetic.

Figure 3.A.13: Replication of Redding and Weinstein (2020) with Nielsen Consumer Panel



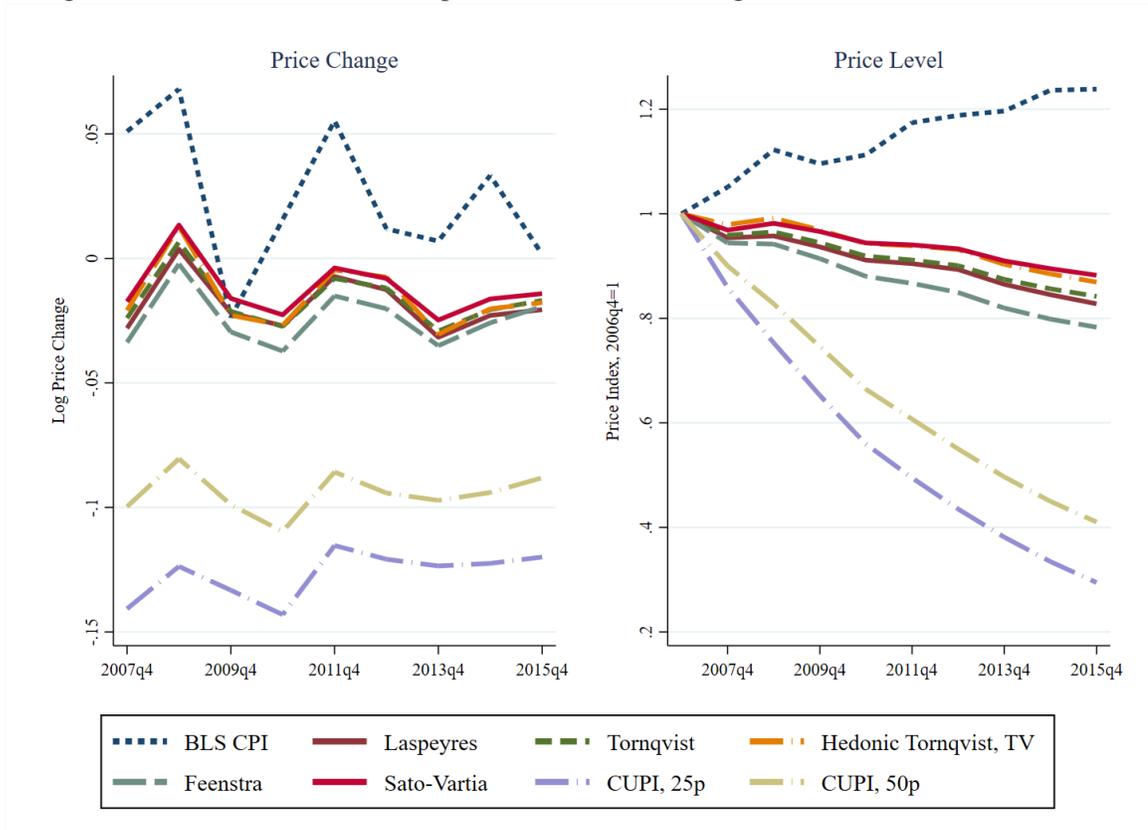
Notes: Figure uses Nielsen Consumer Panel for food and nonfood product groups. The indices are YoY for Q4. The series RW(2020) uses the same CGR duration rule as in Redding and Weinstein (2020). The series 5p(5q) use percentiles based on sales pooled over 5 quarter horizon (current and prior 4 quarters).

Figure 3.A.14: Common Goods Rules – Nielsen Scanner Panel



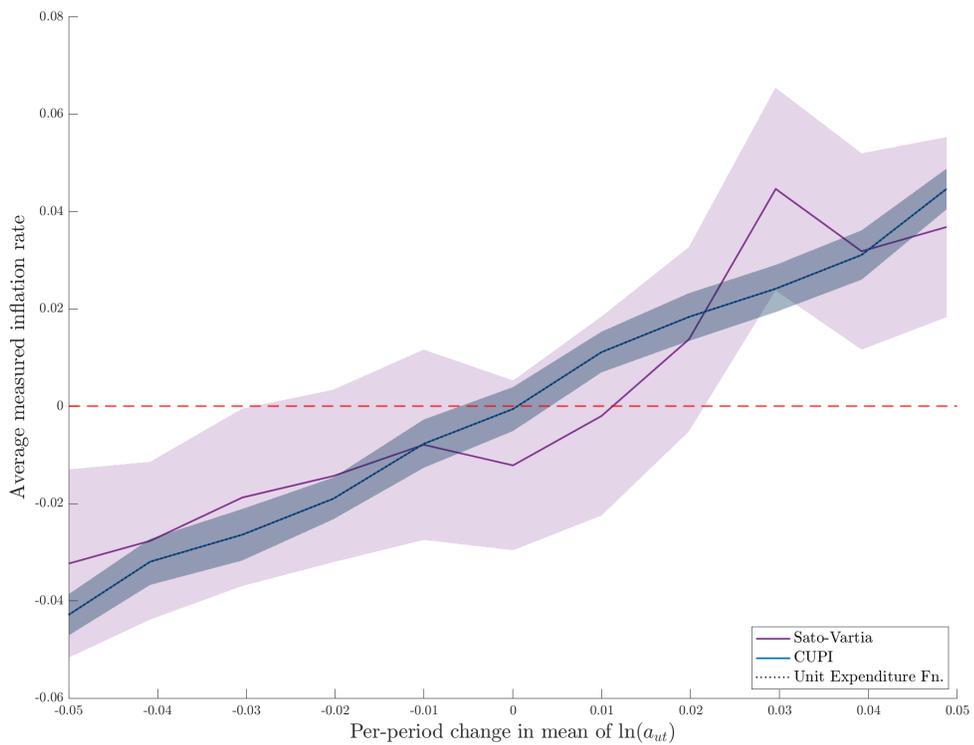
Notes: Figure uses Nielsen Retail Scanner data for food and nonfood product groups. The “CUPI, 25p” and “CUPI, 50p” series use 25th- and 50th-percentile cutoffs for the common goods rule, respectively. The series “CUPI, RW CP” uses the CGR 5th percentile threshold from the consumer Panel data for the common goods rule. Percentiles based on sales pooled over 5 quarter horizon (current and prior 4 quarters). Laspeyres is arithmetic.

Figure 3.A.15: Main Price Index Specifications: Price Changes and Levels, Nielsen Nonfood



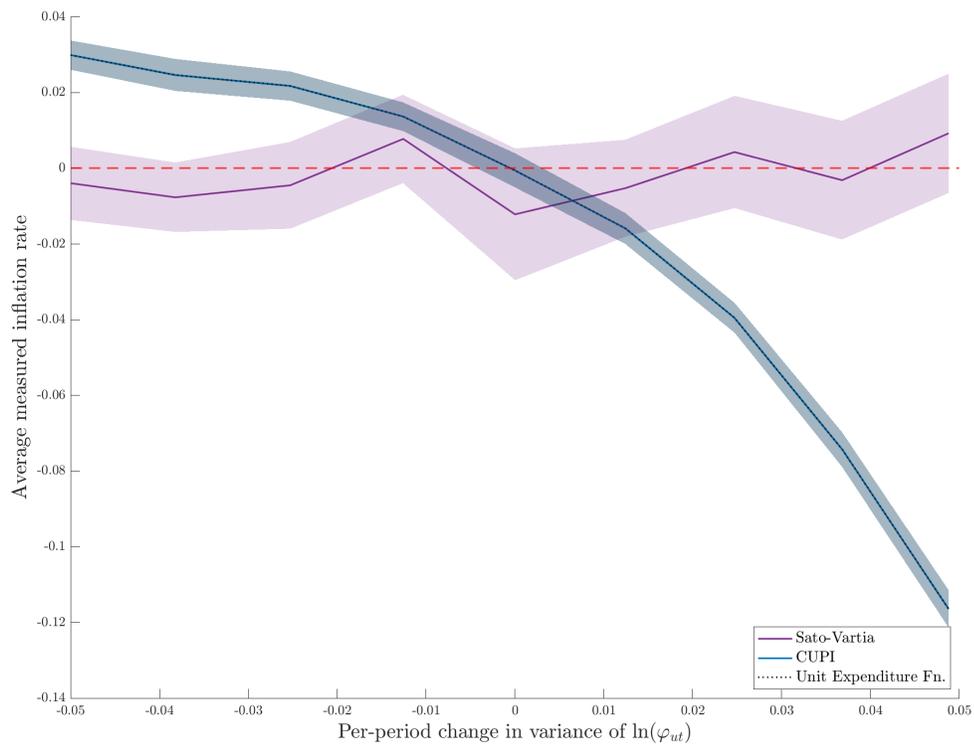
Notes: The figure shows Nielsen Retail Scanner data for nonfood product groups. Price changes show annual log differences from the fourth quarter of the previous year to the fourth quarter of the labeled year. The values are cumulative changes from chained quarterly indices. The price levels chained quarterly values of each price index in the fourth quarter of each year, with the price level in the fourth quarter of 2006 normalized to one for each index. The Laspeyres index is geometric.

Figure 3.A.16: Simulated CES Exact Price Indices with Trends in Cost Shifters



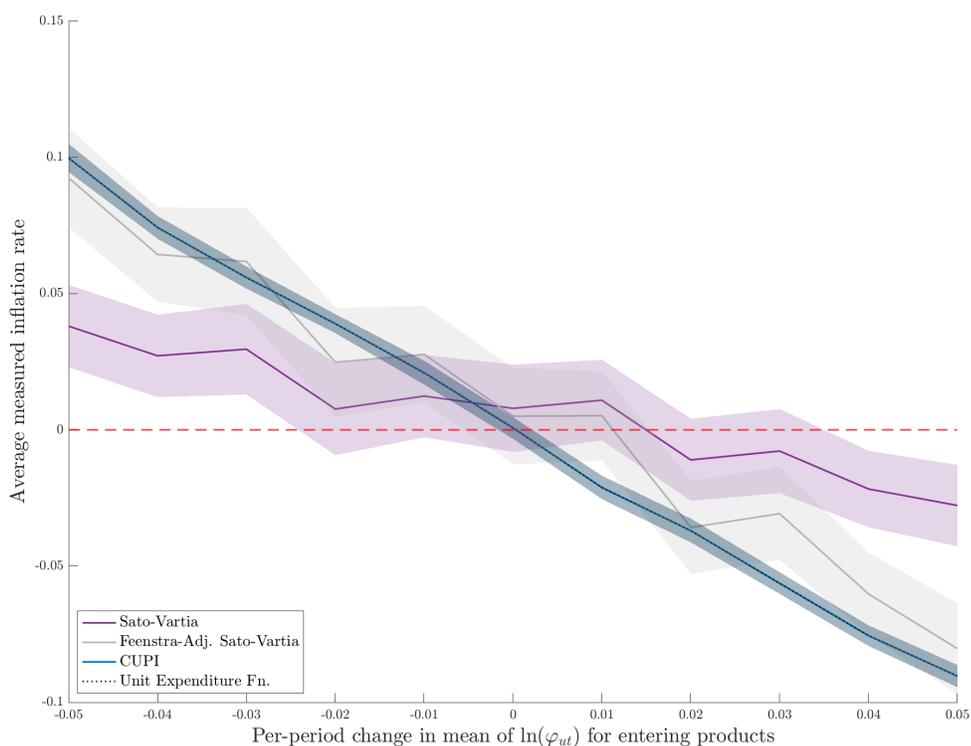
Notes: The figure displays inflation as measured by various CES exact price indices from Monte Carlo simulations of the general equilibrium environment of [Hottman et al. \(2016\)](#). See Section 3.9.3 for simulation details. The horizontal axis displays different average growth rates for the products' marginal cost shifters. The vertical axis displays the average log inflation rate across simulation periods; solid lines represent simple averages across simulations and shaded regions represent 95-percent asymptotic confidence intervals. The CUPI coincides exactly with the unit expenditure function in these simulations.

Figure 3.A.17: Simulated CES Exact Price Indices with Trends in Dispersion of Product Appeal



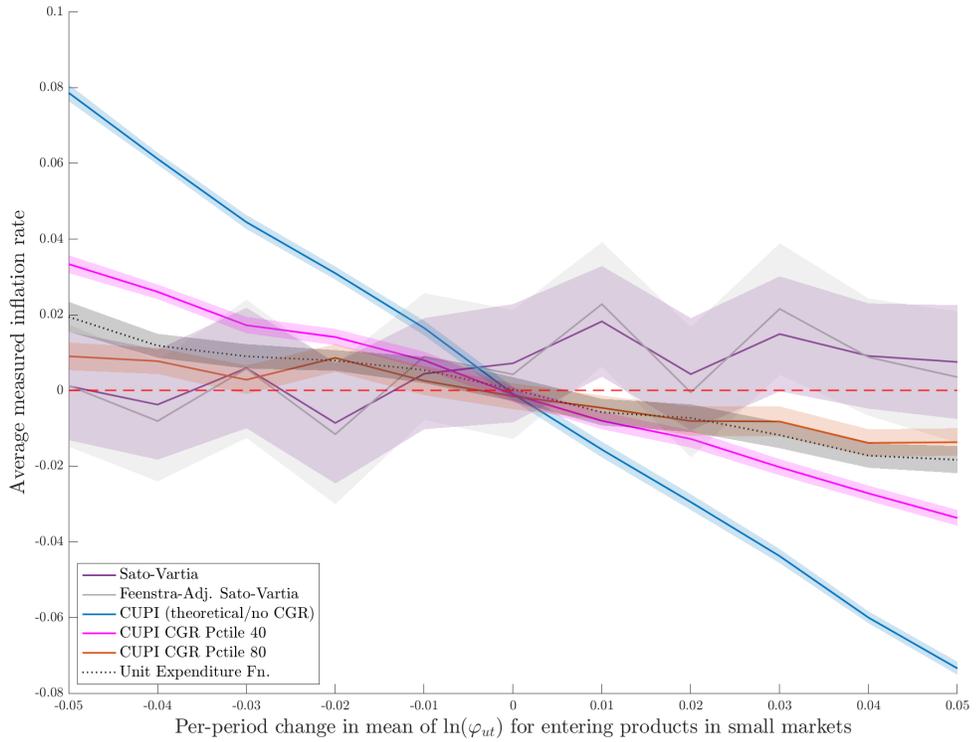
Notes: The figure displays inflation as measured by various CES exact price indices from Monte Carlo simulations of the general equilibrium environment of [Hottman et al. \(2016\)](#). See Section 3.9.3 for simulation details. The horizontal axis displays different average growth rates for the variance of the product appeal parameters. The vertical axis displays the average log inflation rate across simulation periods; solid lines represent simple averages across simulations and shaded regions represent 95-percent asymptotic confidence intervals. The CUPI coincides exactly with the unit expenditure function in these simulations.

Figure 3.A.18: Simulated CES Exact Price Indices with Product Turnover



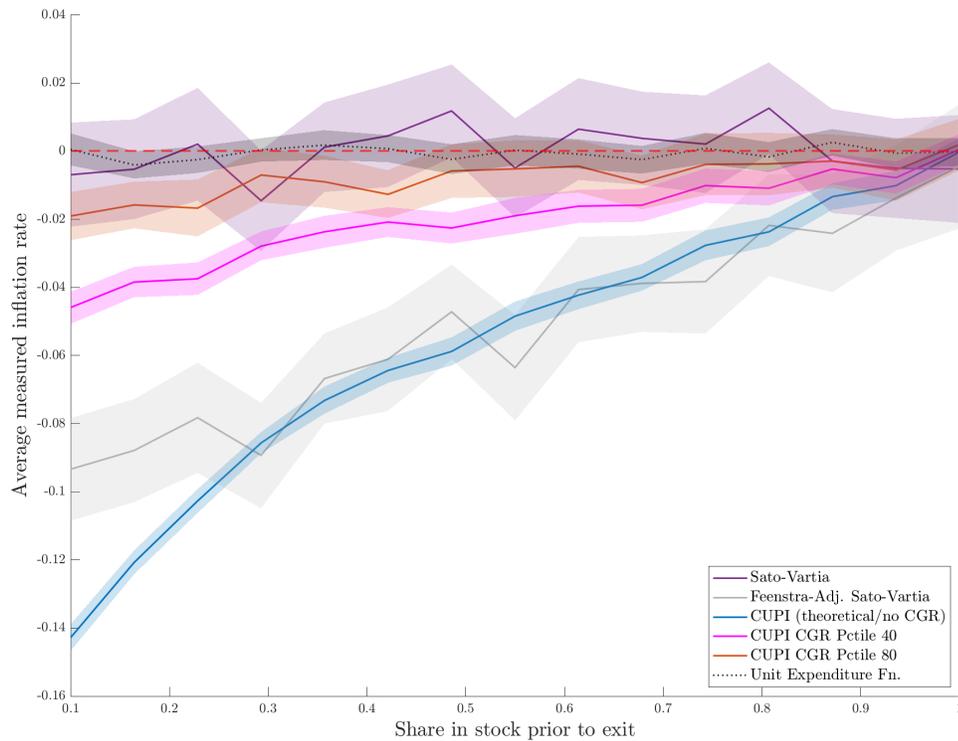
Notes: The figure displays inflation as measured by various CES exact price indices from Monte Carlo simulations of the general equilibrium environment of [Hottman et al. \(2016\)](#). The simulations feature product turnover, with equal numbers of products entering and exiting the market each period. Each product spends five periods in the market. See Section 3.9.3 for simulation details. The horizontal axis displays different average growth rates for the product appeal parameters of entering products. The vertical axis displays the average log inflation rate across simulation periods; solid lines represent simple averages across simulations and shaded regions represent 95-percent asymptotic confidence intervals. The CUPI coincides exactly with the unit expenditure function in these simulations.

Figure 3.A.19: Simulated CES Exact Price Indices with Segmented Markets



Notes: The figure displays inflation as measured by various CES exact price indices from Monte Carlo simulations of the general equilibrium environment of [Hottman et al. \(2016\)](#). The simulations feature segmented markets, with one large “national” market and four small “local” markets. See Section 3.9.3 for simulation details. The horizontal axis displays different average growth rates for the product appeal parameters of entering products in the small markets. The vertical axis displays the average log inflation rate across simulation periods; solid lines represent simple averages across simulations and shaded regions represent 95-percent asymptotic confidence intervals.

Figure 3.A.20: Simulated CES Exact Price Indices with Partial Stock-outs prior to Exit



Notes: The figure displays inflation as measured by various CES exact price indices from Monte Carlo simulations of the general equilibrium environment of [Hottman et al. \(2016\)](#). The simulations feature partial stock-outs in the period prior to products' exit. See Section 3.9.3 for simulation details. The horizontal axis displays the share of the desired quantities available for purchase in the period prior to exit. The vertical axis displays the average log inflation rate across simulation periods; solid lines represent simple averages across simulations and shaded regions represent 95-percent asymptotic confidence intervals.

Chapter 4: Conclusion

This dissertation has presented original research into the causes and consequences of geographic labor mobility with respect to US labor markets, and improvements in the measurement of quality-adjusted inflation.

Chapters 1 and 2 present new descriptive statistics on the rate of job switching within and between US metropolitan areas, showing that non-local job-to-job flows are surprisingly common. Chapter 1 analyses the microeconomic implications of geographic labor mobility for workers through the lens of on-the-job search and wage bargaining, hypothesising that increases in nonlocal labor demand lead to greater job-to-job outflows and wage growth for job stayers. I find evidence in support of both of these claims. The effects of nonlocal labor demand shocks are economically significant, at about 30-50% of similarly estimated effects of local labor demand on local job switching and wage growth. Chapter 2 turns to the macroeconomic implications of geographic labor mobility. I develop a general equilibrium model of spatial, on-the-job search with explicit spatial frictions that inhibit geographic labor mobility. The model allows me to explore the counterfactual effects of changes in these underlying frictions on employment, productivity, and earnings, taking into account the endogeneity of employer market power through wage bargaining and workers' outside options. I find that about half of the wage gains induced by relaxing geographic frictions are due to reduced employer market power. There is room for future research on the micro and macro implications of geographic labor mobility. On the microeconomic side, the analysis of the heterogeneous effects of non-local labor demand shocks is limited to education and industry, largely due to sample size restrictions in the data. Future research can explore heterogeneity in greater detail by initial income, at a more detailed industry level, and potentially by occupation. My results also do not speak to the contribution of non-local

job switching to lifetime income growth, another fruitful area for future research. On the macroeconomic side, the model and results in Chapter 2 describe the long-run effects of changes in geographic frictions and improved labor mobility, but the implications of these same frictions for business cycles and regional adjustment are equally as important. The evidence of heterogeneous exposure to non-local labor demand shocks presented in Chapter 1 provides further motivation for the study of the transmission of labor market shocks across regions with heterogeneous labor market connections.

Chapter 3 presents research on the improvement of inflation measurement with methods that can account for product turnover and heterogeneous and changing product quality. This research is representative of an ongoing body of work with the co-authors of this chapter and other researchers into the improvement of national statistics using high-frequency, micro-level retail sales data. Chapter 3 presents estimates of hedonic price indices that have been constructed with manual, researcher input - a method that is less feasible in the production of timely, large-scale inflation measurement with evolving product markets. Ongoing research is exploring methods for automated model selection for hedonic indices, including elastic net regression and neural networks, that allow the methods presented in this dissertation to scale at a low cost. This will enable more detailed comparisons with official statistics on inflation and nominal sales. In addition, there are several open questions on the proper use of high-frequency detailed scanner data that ongoing work is addressing. This includes methods to evaluate and mitigate chain drift, or the bias induced by non-transitivity of share-weighted index numbers, which is a particular concern with high-frequency measurement of inflation; as well as research into alternative statistical targets that may be more appropriate for elementary price indexes such as a quality-adjusted unit value index.

Bibliography

- R. Adão, C. Arkolakis, and E. Federico. Spatial Linkages, Global Shocks, and Local Labor Markets: Theory and Evidence. 2018.
- D. Arnold. Mergers and Acquisitions , Local Labor Market Concentration , and Worker Outcomes. 2019.
- E. Artuç and J. McLaren. Trade policy and wage inequality: A structural analysis with occupational and sectoral mobility. *Journal of International Economics*, 97(2):278–294, 2015. ISSN 18730353. doi: 10.1016/j.jinteco.2015.06.001. URL <http://dx.doi.org/10.1016/j.jinteco.2015.06.001>.
- E. Artuc, S. Chaudhuri, and J. McLaren. Trade Shocks and Labor Adjustment: Theory. *National Bureau of Economic Research Working Paper Series*, No. 13463(June 2007), 2007. URL <http://www.nber.org/papers/w13463><http://www.nber.org/papers/w13463.pdf>.
- D. H. Autor, D. Dorn, and G. H. Hanson. The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, 103(6):2121–2168, 2013. ISSN 00028282. doi: 10.1257/aer.103.6.2121.
- J. Azar, I. E. Marinescu, and M. Steinbaum. Labor Market Concentration. 2017.
- J. Azar, S. T. Berry, and I. E. Marinescu. Estimating Labor Market Power. *SSRN Electronic Journal*, 2019. doi: 10.2139/ssrn.3456277.
- P. Bajari, Z. Cen, V. Chernozhukov, M. Manukonda, J. Wang, R. Huerta, J. Li, L. Leng, G. George Monokroussos, S. Vijaykumar, and S. Wan. Hedonic prices and quality adjusted price indices powered by ai. Technical report, The Institute for Fiscal Studies Department of Economics, UCL, CENMAP working paper CWP 04/21, 2021.
- W. A. Barnett and K.-H. Choi. Operational identification of the complete class of su-

- relative index numbers: An application of Galois theory. *Journal of Mathematical Economics*, 44(7):603–612, 2008. ISSN 0304-4068. doi: <https://doi.org/10.1016/j.jmateco.2006.05.002>. URL <https://www.sciencedirect.com/science/article/pii/S0304406806000589>. Special Issue in Economic Theory in honor of Charalambos D. Aliprantis.
- A. W. Bartik. Worker Adjustment to Changes in Labor Demand : Evidence from Longitudinal Census Data. *Mimeo*, 2017.
- T. Bartik. *Who Benefits from State and Local Economic Development Policies?* 1991.
- T. J. Bartik. The Effects of Local Labor Demand on Individual Labor Market Outcomes for Different Demographic Groups and the Poor. *Journal of Urban Economics*, 40(93): 150–178, 1996.
- T. J. Bartik. How Effects of Local Labor Demand Shocks Vary with the Initial Local Unemployment Rate. 2015.
- P. Beaudry, D. A. Green, and B. M. Sand. Spatial equilibrium with unemployment and wage bargaining: Theory and estimation. *Journal of Urban Economics*, 79:2–19, 2014. ISSN 00941190. doi: 10.1016/j.jue.2013.08.005.
- M. Beine, S. Bertoli, J. Fernández, and H. Moraga. A practitioners’ guide to gravity models of international migration. *The World Economy*, 29(4), 2016.
- C. L. Benkard and P. Bajari. Hedonic price indexes with unobserved product characteristics, and application to personal computers. *Journal of Business and Economic Statistics*, 23 (1):61–75, jan 2005. ISSN 07350015. doi: 10.1198/073500104000000262.
- E. Benmelech, N. Bergman, and H. Kim. Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages? 2018.
- D. Berger, K. Herkenhoff, and S. Mongey. Labor Market Power. *American Economic Review*, 112(4), 2022.
- A. Bilal. The Geography of Unemployment *. Technical report, 2019. URL <https://sites.google.com/site/adrienbilal>.
- O. J. Blanchard and L. F. Katz. Regional evolutions. *Brookings Papers on Economic Activity*, 1992(1):1–75, 1992. ISSN 15334465. doi: 10.2307/2534556.
- K. Borusyak, P. Hull, and X. Jaravel. Quasi-Experimental Shift-Share Research Designs. 2020.

- C. Broda and D. Weinstein. Product creation and destruction: Evidence and price implications. *American Economic Review*, 100:691–723, 2010.
- K. Burdett and D. Mortensen. Wage Differentials, Employer Size, and Unemployment. *International Economic Review*, 1998.
- Bureau of Labor Statistics. Quality adjustment in the cpi, 2023. URL <https://www.bls.gov/cpi/quality-adjustment/home.htm>. Accessed February 9, 2023.
- D. M. Byrne, D. E. Sichel, and A. Aizcorbe. Getting Smart About Phones: New Price Indexes and the Allocation of Spending Between Devices and Services Plans in Personal Consumption Expenditures. *Finance and Economics Discussion Series*, 2019 (012), 2019. ISSN 19362854. doi: 10.17016/feds.2019.012.
- M. Caballero, B. C. Cadena, and B. Kovak. The International Transmission of Local Economic Shocks Through Migrant Networks. *SSRN Electronic Journal*, 2021. doi: 10.2139/ssrn.3829394.
- M. J. Cafarella, G. Ehrlich, T. Gao, M. D. Shapiro, and L. Y. Zhao. Using machine learning to construct hedonic price indices. Unpublished working paper, Census, Maryland and Michigan, 2023.
- P. Cahuc, F. Postel-Vinay, and J. M. Robin. Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, 74(2):323–364, 2006. ISSN 00129682. doi: 10.1111/j.1468-0262.2006.00665.x.
- S. Caldwell and O. Danieli. Outside Options in the Labor Market. 2018.
- S. Caldwell and N. Harmon. Outside Options, Bargaining, and Wages: Evidence from Coworker Networks. 2019.
- D. Card, U. C. Berkeley, J. Rothstein, and M. Yi. Location, Location, Location. 2021.
- G. Chodorow-Reich and J. Wieland. Secular labor reallocation and business cycles. *Journal of Political Economy*, 128(6):2245–2287, 2020. ISSN 1537534X. doi: 10.1086/705717.
- M. Dao, D. Furceri, and P. Loungani. Regional Labor Market Adjustment in the United States: Trend and Cycle. *The Review of Economics and Statistics*, 99(May):243–257, 2017. doi: 10.1162/REST.
- W. Dauth, S. Findeisen, E. Moretti, and J. Suedekum. Matching in Cities. 2018. URL <http://www.nber.org/papers/w25227>.

- S. J. Davis and J. C. Haltiwanger. Labor Market Flows in the Cross-Section and Over Time. 2011.
- J. De Haan. Hedonic price indexes: a comparison of imputation, time dummy and other approaches. 2008.
- J. De Haan and H. A. Van Der Grient. Eliminating chain drift in price indexes based on scanner data. *Journal of Econometrics*, 161(1):36–46, mar 2011. doi: 10.1016/j.jeconom.2010.09.004.
- E. Diewert. Quality adjustment and hedonics: A unified approach. Working paper, University of British Columbia, 2019.
- E. Diewert. Chapter 6: The Chain Drift Problem and Multilateral Indices. 2021.
- E. Diewert, S. Heravi, and M. Silver. Hedonic Imputation versus Hedonic Time Dummy Indexes. Technical report, NBER WP 14018, 2008.
- W. E. Diewert. Superlative index numbers and consistency in aggregation. *Econometrica: Journal of the Econometric Society*, pages 883–900, 1978.
- F. Eckert, S. Ganapati, and C. Walsh. Skilled Scalable Services: The New Urban Bias in Recent Economic Growth. *CESifo Working Paper*, 2020.
- G. Ehrlich, J. Haltiwanger, R. Jarmin, D. Johnson, and M. D. Shapiro. Re-engineering key national economic indicators. *Big Data for Twenty-First Century Economic Statistics*, 2022.
- O. Eltető and P. Köves. On a problem of index number computation relating to international comparison. *Statisztikai Szemle*, 42(10):507–518, 1964. (in Hungarian).
- T. Erickson and A. Pakes. An experimental component index for the CPI: From annual computer data to monthly data on other goods. *American Economic Review*, 101(5): 1707–1738, aug 2011. ISSN 00028282. doi: 10.1257/aer.101.5.1707.
- R. C. Feenstra. New Product Varieties and the Measurement of International Prices. *The American Economic Review*, 84(1):157–177, 1994.
- R. C. Feenstra and M. B. Reinsdorf. *Should Exact Index Numbers Have Standard Errors? Theory and Application to Asian Growth*. Number October. 2007. ISBN 9780226044491. doi: 10.3386/w10197. URL <http://www.nber.org/papers/w10197>{%}5Cn<http://www.nber.org/papers/w10197.pdf>.

- J. Flemming. Costly Commuting and the Job Ladder. *Finance and Economics Discussion Series*, 2020(025), mar 2020. ISSN 19362854. doi: 10.17016/FEDS.2020.025.
- C. Gini. On the circular test of index numbers. *Metron*, 9(9):3–24, 1931.
- C. F. Goetz. Unemployment Duration and Geographic Mobility: Do Movers Fare Better than Stayers? *SSRN Electronic Journal*, 2015. doi: 10.2139/ssrn.2573056.
- P. Goldsmith-Pinkham, I. Sorkin, and H. Swift. Bartik Instruments : What, When, Why, and How. *American Economic Review*, 110(8):2586–2624, 2020.
- A. Gorajek. Generalizing the stochastic approach to price indices. *Review of Income and Wealth*, pages 1–22, October 2022.
- J. Haltiwanger, H. Hyatt, and E. McEntarfer. Do Workers Move Up the Firm Productivity Job Ladder? (November):1–47, 2017.
- J. Haltiwanger, H. Hyatt, and E. McEntarfer. Who moves up the job ladder? *Journal of Labor Economics*, 36(S1):S301–S336, 2018. ISSN 0734306X. doi: 10.1086/694417.
- K. Head and T. Mayer. *Gravity Equations: Workhorse, Toolkit, and Cookbook*, volume 4. Elsevier B.V., 2014. ISBN 9780444543141. doi: 10.1016/B978-0-444-54314-1.00003-3. URL <http://dx.doi.org/10.1016/B978-0-444-54314-1.00003-3>.
- S. Heise and T. Porzio. Spatial Wage Gaps and Frictional Labor Markets. Technical report, 2019.
- R. J. Hill and D. Melser. The Hedonic Imputation Method and the Price Index Problem. Technical report, 2006.
- C. J. Hottman, S. J. Redding, and D. E. Weinstein. Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364, 2016.
- C.-T. Hsieh and E. Moretti. Housing Constraints and Spatial Misallocation. 2015.
- H. Hyatt, K. Mckinney, E. Mcentarfer, S. Tibbets, L. Vilhuber, and D. Walton. Job-to-Job Flows: New Statistics on Worker Reallocation and Job Turnover. May:1–25, 2015.
- H. Hyatt, E. Mcentarfer, K. Ueda, and A. Zhang. Interstate Migration and Employer-to-Employer Transitions in the U.S.: New Evidence from Administrative Records Data. 2018.
- H. R. Hyatt and J. Spletzer. The Recent Decline in Employment Dynamics. 2013.

- L. Ivancic, W. E. Diewert, and K. J. Fox. Scanner data, time aggregation and the construction of price indexes. *Journal of Econometrics*, 161(1):24–35, 2011.
- D. A. Jaeger, J. Stuhler, M. M.-c. S, H. We, J. Angrist, M. Amior, G. Borjas, C. Dustmann, A. Edo, J. F.-h. Moraga, J. Hunt, L. Katz, J. Llull, M. Manacorda, S. Markussen, E. Murard, B. Petrongolo, U. Schönberg, J. C. S. Serrato, and U. Sunde. Shift-Share Instruments and the Impact of Immigration. 2018.
- S. Jager, B. Schoefer, S. Young, and J. Zweimuller. Wages and the Value of Nonemployment. 2019.
- G. Kaplan and S. Schulhofer-Wohl. Understanding the Long-Run Decline in Interstate Migration. *International Economic Review*, 58(1):57–94, 2017. ISSN 14682354. doi: 10.1111/iere.12209.
- F. Karahan and S. Rhee. Population Aging, Migration Spillovers, and the Decline in Interstate Migration. 2014.
- J. Kennan and J. Walker. The Effect of Expected Income on Individual Migration Decisions. *Econometrica*, 2011. doi: 10.3982/ecta4657.
- A. A. Konüs. The problem of the true index of the cost of living. *Econometrica*, pages 10–29, 1939.
- G. Kosar, T. Ransom, and H. W. van der Klaauw. Understanding Migration Aversion Using Elicited Counterfactual Choice Probabilities. *SSRN Electronic Journal*, 2019. doi: 10.2139/ssrn.3368408.
- M. Lachowska, A. Mas, R. Saggio, and S. A. Woodbury. Do Workers Bargain over Wages? A Test Using Dual Jobholders. 2021.
- A. Manning and B. Petrongolo. How local are labor markets? Evidence from a spatial job search model. *American Economic Review*, 107(10):2877–2907, 2017. ISSN 00028282. doi: 10.1257/aer.20131026.
- I. Marinescu and R. Rathelot. Mismatch Unemployment and the Geography of Job Search. 2018.
- I. Marinescu, I. Ouss, and L.-D. Pape. Wages, Hires, and Labor Market Concentration. 2019.
- R. Martin. Changing Tastes Versus Specification Error in Cost-of-Living Measurement. BLS Working Paper 531, 2020.

- C. Meghir, R. Narita, and J. M. Robin. Wages and informality in developing countries. *American Economic Review*, 105(4):1509–1546, 2015. ISSN 00028282. doi: 10.1257/aer.20121110.
- A. Mian, K. Rao, and A. Sufi. Household Balance Sheets, Consumption, and the Economic Slump. *Quarterly Journal of Economics*, 128(4):1687–1726, 2013.
- R. Molloy and C. Smith. U.S. Internal Migration: Recent Patterns and Outstanding Puzzles. 2019.
- R. Molloy, C. L. Smith, and A. Wozniak. Declining Migration Within the US: The Role of the Labor Market. 2014.
- R. Molloy, C. L. Smith, R. Trezzi, and A. Wozniak. Understanding Declining Fluidity in the U.S. Labor Market. *Brookings Papers on Economic Activity Conference Draft*, 2016.
- R. Molloy, C. L. Smith, and A. Wozniak. Job Changing and the Decline in Long-Distance Migration in the United States. *Demography*, 54(2):631–653, 2017. ISSN 15337790. doi: 10.1007/s13524-017-0551-9.
- F. Monte, S. J. Redding, and E. Rossi-Hansberg. Commuting, Migration, and Local Employment Elasticities. *American Economic Review*, 108(12):3855–3890, 2018. ISSN 0002-8282. doi: 10.1257/aer.20151507.
- E. Moretti. Local labor markets. *Handbook of Labor Economics*, 4b:1237–1313, 2010. doi: 10.4324/9780203828892.
- G. Moscarini and F. Postel-Vinay. Stochastic Search Equilibrium. *Review of Economic Studies*, 2013.
- National Retail Federation. 4-5-4 calendar, 2023. URL <https://nrf.com/resources/4-5-4-calendar#category.name>. Accessed February 10, 2023.
- M. J. Notowidigdo. The incidence of local labor demand shocks. *Journal of Labor Economics*, 38(3):687–725, 2020. ISSN 0734306X. doi: 10.1086/706048.
- A. Pakes. A reconsideration of hedonic price indexes with an application to PC’s. *American Economic Review*, 93(5):1578–1596, 2003. ISSN 00028282. doi: 10.1257/000282803322655455.
- T. Ransom. Labor Market Frictions and Moving Costs of the Employed and Unemployed. 2018. URL <https://tyleransom.github.io/research/movingCostsLMfrictions.pdf>.

- S. Redding and D. E. Weinstein. Measuring aggregate price indexes with taste shocks: Theory and evidence for ces preferences. *Quarterly Journal of Economics*, 135:503–560, 2018.
- S. J. Redding and D. E. Weinstein. Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences. *The Quarterly Journal of Economics*, 135 (1):503–560, 2020. ISSN 0033-5533. doi: 10.1093/qje/qjz031.
- K. Rinz. Labor Market Concentration, Earnings Inequality, and Earnings Mobility. 2018.
- J. Roback. Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90(6): 1257–1278, 1982.
- K. Sato. The Ideal Log-Change Index Number. *The Review of Economics and Statistics*, 1976. ISSN 00346535. doi: 10.2307/1924029.
- C. Schluter and G. Willems. A Dynamic Empirical Model of Frictional Spatial Job Search *. Technical report, 2018.
- B. Schmutz and M. Sidibé. Frictional Labour Mobility. *The Review of Economic Studies*, sep 2018. ISSN 0034-6527. doi: 10.1093/restud/rdy056.
- G. Schubert. House Price Contagion and U . S . City Migration Networks. 2021.
- G. Schubert, A. Stansbury, and B. Taska. Monopsony and Outside Options. 2020.
- H. Schultz. A Misunderstanding in Index-Number Theory: The True Konus Condition on Cost-of-Living Index Numbers and Its Limitations. *Econometrica*, 1939. doi: 10.2307/1906996.
- B. Sculz. Index numbers for multiregional comparisons. *Przegląd Statystyczny*, 3:239–254, 1964.
- R. Shimer. On-the-job search and strategic bargaining. *European Economic Review*, 50(4): 811–830, 2006. ISSN 00142921. doi: 10.1016/j.eurocorev.2006.01.008.
- R. H. Topel and M. P. Ward. Job Mobility and the Careers of Young Men. *Quarterly Journal of Economics*, 107(2):439–479, 1992.
- Y. Vartia. Ideal Log-Change Index Numbers. *Scandinavian Journal of Statistics*, 3(3): 121–126, 1976.
- W. Zhang. Distributional Effects of Local Minimum Wage Hikes: A Spatial Job Search

Approach. Technical report, 2017. URL <http://laborcenter.berkeley.edu/minimum-wage-living->.

A. Şahin, J. Song, G. Topa, and G. L. Violante. Mismatch unemployment. *American Economic Review*, 104(11):3529–3564, 2014. ISSN 00028282. doi: 10.1257/aer.104.11.3529.