Income volatility, taxation and the functioning of the U.S. labor market*

Thibaut Lamadon[†] Magne Mogstad[‡] Bradley Setzler[§]

This Version: April 19, 2021

Abstract

The goal of this report is to characterize income volatility in the U.S. labor market and examine its causes and consequences. To achieve this goal, we analyze U.S. business and household tax records. These administrative data sets allow us to match employees and employers and to construct panel data on the outcomes and characteristics of U.S. firms, individuals and households. The main insights from the empirical analysis may be summarized in four broad conclusions. First, income volatility rose steadily during 2001-2009, peaked during the Great Recession, then dropped during 2010-2015. However, these national averages miss a lot. Income volatility is relatively high on the coasts and in the Western part of the country, and lower socioeconomic areas tend to have higher income volatility. Second, income volatility is lower if one considers income net of taxes and transfers. In particular, the Federal tax-transfer system attenuates both permanent shocks at the worker level and the pass-through of firm shocks to workers' earnings. Third, worker mobility across firms generates relatively small changes in income. By contrast, sorting of better workers to better firms and the exit and entry of firms in local markets are empirically important determinants of workers' income. Fourth, income volatility, at the individual and market level, may generate substantial changes in tax payments and the receipt of tax credits. This indicates that income volatility in the U.S. labor market could make it difficult to obtain accurate predictions of tax revenues.

*The opinions expressed in this report are those of the authors alone and do not reflect the views of the Internal Revenue Service or the U.S. Treasury Department. This work is a component of a larger project on income risk in the United States, conducted through the SOI Joint Statistical Research Program. We are grateful to Victoria Bryant, Raj Chetty, Nathan Hendren, Kevin Pierce, Michael Weber, Danny Yagan, Owen Zidar, and Eric Zwick for support and guidance in accessing and understanding the IRS data. We also appreciate the constructive comments and suggestions from discussants and participants at various conferences and seminars. Mogstad and Setzler acknowledge funding from the Washington Center for Equitable Growth.

[†]Department of Economics, University of Chicago. E-mail: lamadon@uchicago.edu

[‡]Department of Economics, University of Chicago; Statistics Norway; NBER; IFS. E-mail: magne.mogstad@gmail.com

[§]Department of Economics, University of Chicago. E-mail: bradley.setzler@gmail.com

1 Introduction

The aim of this report is to characterize income volatility in the U.S. labor market and examine its causes and consequences. There are a number of key questions addressed. What is the distribution of income volatility in the U.S. labor market? How large and persistent are the year-by-year changes in the incomes of American workers? How much does the sorting of workers to firms, regions, and industries matter for income volatility? How do the estimates of income volatility change when we account for other sources of income such as spousal earnings? To what extent does the Federal tax-and-tranfer system affect measures of income volatility? Does income volatility make it difficult to predict tax revenues or receipt of tax credits?

Data challenges have made it difficult in the past to answer these questions. The ideal data covers a large number of individuals, includes a sufficient number of years on each individual, links each individual to her employer, and links individuals to households. While such data has not previously been available for the U.S., administrative data has provided such information for existing studies on some other countries. The advantages of these administrative data sets are the accuracy of the income information provided, the large sample size, and the lack of attrition, other than what is due to migration and death, as well as the possibility to link to employers and households.

To investigate the above questions, we analyze U.S. business and household tax records. These administrative data sets allow us to match employees and employers and to construct panel data on the outcomes and characteristics of U.S. firms, individuals and households. The main insights from the empirical analysis may be summarized in four broad conclusions. First, income volatility rose steadily during 2001-2009, peaked during the Great Recession, then dropped during 2010-2015. However, these national averages miss a lot. Income volatility is relatively high on the coasts and in the Western part of the country, and lower socioeconomic areas tend to have higher income volatility. Second, income volatility is lower if one considers income net of taxes and transfers. In particular, the Federal tax-transfer system attenuates both permanent shocks at the worker level and the pass-through of firm shocks to workers' earnings. Third, worker mobility across firms generate relatively small changes in income. By contrast, sorting of better workers to better firms and the exit and entry of firms in local markets are empirically important determinants of workers' income. Fourth, income volatility, at the individual and market level, may generate substantial changes in tax payments and the receipt of tax credits. This indicates that income volatility in the U.S. labor market could make it difficult to obtain accurate predictions of tax revenues.

Our work relates to a considerable literature on income volatility, risk, and inequality.[1] DeBacker et al. (2013) use a panel of tax returns to study the persistent-versus-transitory nature of rising inequality in individual male labor earnings and in total household income, both before and after taxes, in the U.S. Their paper is the first to estimate error components models of income dynamics using U.S. administrative data.[2] Building on this work, we characterize the

[1]See, for example, the recent review by Meghir and Pistaferri (2011), and the extensive list of studies referenced therein.

[2]See also Blundell et al. (2015) who perform a similar analysis for Norway.

variation over time and across areas in income volatility in the U.S. and explore the factors correlated with high income volatility. Moreover, we separate between income volatility caused by idiosyncratic shocks to individual workers and the volatility reflecting firm shocks common to workers in that firm. We also explore how the sorting of workers to firms, regions, and industries matters for the volatility and inequality in income. Our report also adds to existing work in that we compare volatility in earnings, household gross income and household net income. This allows us to draw inference about how the family and the tax-transfer system attenuate income volatility.

Our analysis also contributes to a large and growing literature on firms, income volatility and labor market inequality, reviewed in Card et al. (2018). A number of studies show that trends in wage dispersion closely track trends in productivity dispersion across industries and workplaces (Faggio et al., 2010; Dunne et al., 2004; Barth et al., 2016). While this correlation might reflect that some of the productivity differences across firms spill over to wages, it could also be driven by changes in the degree to which workers of different quality sort into different firms (see e.g. Murphy and Topel, 1990; Gibbons and Katz, 1992; Gibbons et al., 2005). To address the sorting issue, a growing body of work has taken advantage of matched employer-employee data. Some studies use this data to estimate the pass-through of changes in the value added of a firm to the wages of its workers, while controlling for time-invariant firm and worker heterogeneity (see e.g. Guiso et al., 2005; Card et al., 2013a; Card et al., 2018; Carlsson et al., 2016; Balke and Lamadon, 2020; Friedrich et al., 2019). These studies typically report estimates of pass-through in the range of 0.05-0.20. We complement this work by providing evidence of pass-through for a broad set of firms in the U.S. and by showing how the estimated pass-through of firm shocks is confounded by market shocks and attenuated by the tax-transfer system.

Another set of studies use the matched employer-employee data to estimate the changes in earnings caused by workers moving across firms. Following Abowd et al. (1999), these studies typically use an additive worker and firm effects model. They tend to conclude that firms play an important role in the determination of earnings, with a typical finding that about 15-20 percent of the variance of log earnings is attributable to the choice of firm (Card et al., 2018). We show, however, that firm effects are small in the U.S. labor market, explaining only a few percent of the variation in earnings. This finding contrasts with recent work from the U.S. (Sorkin, 2018; Song et al., 2018) as well as many studies from other developed countries (Card et al., 2018). The reason is that these studies do not address the concern that estimates of firm effects will be biased upward and estimates of worker sorting will be biased downward in finite samples, with the size of the bias depending inversely on the degree of worker mobility among firms (Andrews et al., 2008). Following recent work by Bonhomme et al. (2019) and Kline et al. (2020), we apply two alternative approaches to correct for the bias of the estimator of Abowd et al. (1999). Both approaches show that firm effects explain very little of the variation in earnings in the U.S. economy, once one corrects for bias due to limited mobility. Instead, a substantial part of the variation in earnings is due to positive sorting of high wage workers to high paying firms. Our report also differs in that we estimate the additive worker and firm effects model both for earnings, household gross income and household net income. This allows us to draw inference about how the progressive nature of the tax-transfer system attenuates the income changes associated with moving across firms and reduces the incentives of better workers to sort into better firms.

The remainder of the report is organized as follows. Section 2 describes the data and the sample selection. Section 3 characterizes income mobility in the U.S. labor market and describe how it varies over time and across areas. In Section 4, we use several complementary approaches to examine causes and consequences of income volatility. Section 5 offers some concluding remarks.

2 Data sources and sample selection

2.1 Data sources

Our empirical analyses are based on a matched employer-employee panel data set with information on the characteristics and outcomes of U.S. workers and firms. This data is constructed by linking U.S. Treasury business tax filings with worker-level filings for the years 2001-2015. Below, we briefly describe data sources, sample selection, and key variables, while details about the data construction and the definition of each of the variables are given in Appendix A.

Business tax returns include balance sheet and other information from Forms 1120 (C-corporations), 1120S (S-corporations), and 1065 (partnerships). The key variables that we draw on from the business tax filings are the firm's value added, commuting zone, and industry code. Value added is the difference between receipts and the cost of goods sold. Commuting zone is constructed using the ZIP code of the firm's business filing address. Industry is defined as the first two digits of the firm's NAICS code. We define a market as the combination of an industry and a commuting zone. At times we will aggregate these markets according to the combination of Census regions (Midwest, Northeast, South, West) and broad sectors (Goods and Services). We will refer to this classification as "broad markets".

Earnings data are based on taxable remuneration for labor services for direct employees and independent contractors. Earnings include wages and salaries, bonuses, tips, exercised stock options, and other sources of income deemed taxable. These forms are filed by the firm on behalf of the worker and provide the firm-worker link. Gross household income is constructed using a definition similar to that of Piketty and Saez (2003). Net household income is given by gross household income minus Federal taxes plus Federal benefits from Social Security and unemployment. See Appendix A for further details.

We express all monetary variables in 2015 dollars, adjusting for inflation using the Consumer Price Index.

2.2 Sample Selection

In each year, we start with all individuals aged 25-60 who are linked to at least one employer. Next, we define the worker's firm as the EIN that pays her the greatest direct (W-2) earnings

	Workers		Firms	
Panel A.	**Baseline Sample**			
	Unique	**Observation-Years**	**Unique**	**Observation-Years**
Full Sample:	89,570,480	447,519,609	6,478,231	39,163,975
Panel B.	**Movers Sample**			
	Unique	**Observation-Years**	**Unique**	**Observation-Years**
Movers Only:	32,070,390	207,990,422	3,559,678	23,321,807
Panel C.	**Stayers Sample**			
	Unique	**6 Year Spells**	**Unique**	**6 Year Spells**
Complete Stayer Spells:	10,311,339	35,123,330	1,549,190	6,533,912
10 Stayers per Firm:	6,297,042	20,354,024	144,412	597,912
10 Firms per Market:	5,217,960	16,506,865	117,698	476,878

Table 1: Overview of the Sample

Notes: This table provides an overview of the full sample, movers sample, and stayers sample, including the steps involved in defining the stayers sample.

in that year. This definition of a firm conforms to previous research using the U.S. business tax records (see, e.g., Song et al., 2018). The EIN defines a corporate unit for tax and accounting purposes. It is a more aggregated concept than an establishment, which is the level of analysis considered in recent research on U.S. Census data (see, e.g., Barth et al., 2016), but a less aggregated concept than a parent corporation. As a robustness check, we investigated the sensitivity of the estimated firm wage premiums to restricting the sample to EINs that appear to have a single primary establishment. These are EINs for which the majority of workers live in the same commuting zone. It is reassuring to find that the estimated firm wage premiums do not materially change when we use this restricted sample.[3]

Since we do not observe hours worked or a direct measure of full-time employment, we follow the literature by including only workers for whom annual earnings are above a minimum threshold (see, e.g., Song et al., 2018). In the baseline specification, this threshold is equal to \$15,000 per year (in 2015 dollars), which is approximately what people would earn if they work full-time at the federal minimum wage. As a robustness check, we investigate the sensitivity of our results to other choices of a minimum earnings threshold. We further restrict the sample to firms with non-missing value added, commuting zone, and industry. The full sample includes 447.5 (39.2) million annual observations on 89.6 (6.5) million unique workers (firms).

In parts of the analysis, we consider two distinct subsamples. The first subsample, which we refer to as the *stayers sample*, restricts the full sample to workers observed with the same employer for eight consecutive years. This restriction is needed to allow for a flexible specification of how the worker's earnings evolve over time. Specifically, we omit the first and last years of these spells (to avoid concerns over workers exiting and entering employment during the year, confounding the measure of annual earnings) and analyze the remaining six-year spells. Furthermore, the stayers sample is restricted to employers that do not change commuting zone

[3]In the baseline sample, the AKM (BLM) estimates of firm effects are around 10 (3) percent. By comparison, the restricted sample gives AKM (BLM) estimates of approximately 9 (3) percent.

or industry during those eight years. Lastly, we restrict the stayers sample to firms with at least 10 such stayers and markets with at least 10 such firms, which helps to ensure sufficient sample size to perform the analyses at both the firm and the market level. The stayers sample includes 35.1 (6.5) million spells on 10.3 (1.5) million unique workers (firms).

The second subsample, which we refer to as the *movers sample*, restricts the full sample to workers observed at multiple firms. That is, it is not the same EIN that pays the worker the greatest direct (W-2) earnings in all years. Following previous work, we also restrict the movers sample to firms with at least two movers. This restriction might help reduce the limited mobility bias. It also makes it easier to directly compare the AKM and BLM estimates of firm effects to those produced by the approach of Kline et al. (2020) (which requires at least two movers per firm). The movers sample includes 32.1 (3.6) million unique workers (firms).

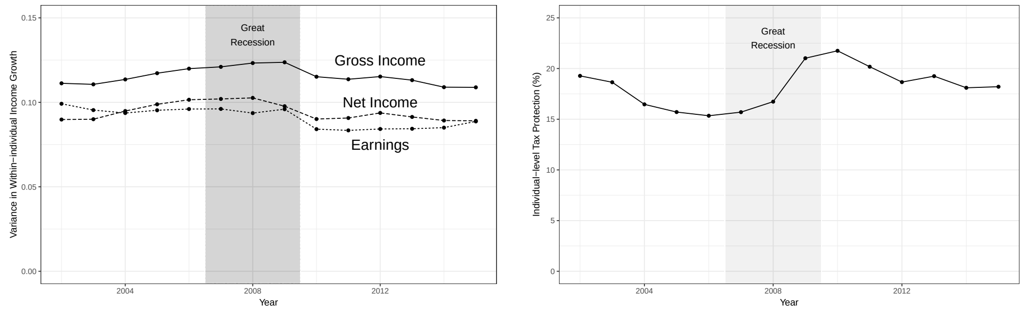
Table 1 compares the size of the baseline, the stayers, and the movers samples. Detailed summary statistics of these samples of linked firms and worker are given in Appendix Table A.1. The samples are broadly similar, both in the distribution of earnings but also in firm-level variables such as value added, wage bill, size, and the geographic distribution across regions and sectors. The most noticeable differences are that the stayers have, on average, somewhat higher earnings and tend to work in firms with higher value added.

3 Income volatility in the U.S.

In this section, we characterize income volatility in the U.S. labor market and describe how it has changed over time and across areas. Following the literature, volatility is defined as movements up or down in a household's income over time, as measured by the variance of the year-by-year changes in log household (annual) income. We construct this measure of volatility for three different measures of income: individual earnings, gross household income, and net household income. Following Blundell et al. (2015), we interpret the reduction in net household income volatility relative to gross household income volatility as a measure of the protection provided by the federal tax-transfer system. The transfers include Social Security benefits, unemployment benefits, and the Earned Income Tax Credit (EITC).

Time trends in income volatility

Figure 1 presents estimates of income volatility across years. In Figure 1(a), we see that gross income volatility rose steadily during 2001-2009, peaking during the Great Recession, then dropped during 2010-2015. Net household income and earnings volatility followed similar trends, but with smaller magnitudes. Figure 1(b) presents one minus the ratio of net household income to gross household income, multiplied by 100%. This is a measure of the reduction in income volatility, or protection, that is due to Federal taxes and transfers. The Federal tax-transfer system provides substantial protection against income volatility. Federal taxes and transfers reduce income volatility by about 21% on average, with a low of around 15% in 2001 and a peak of around 23% in 2009.



(a) Earnings, Gross and Net Household Income Volatility

(b) Percentage of Gross Income Volatility Attenuated by the Federal Tax-Transfer System

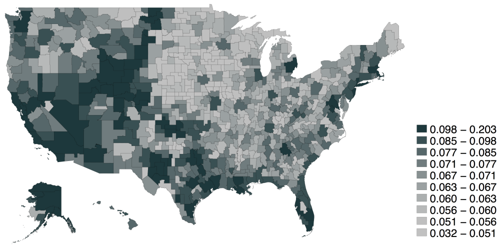
Figure 1: Income Volatility in the United States

Notes: This figure presents (a) the quantity of income volatility in the United States for gross household income, net household income, and earnings, and (b) the percentage reduction in volatility attributable to the federal tax-transfer system. Volatility is defined as movements up or down in a household's income over time, as measured by the variance of the year-by-year changes in log household (annual) income. The volatility reduction due to the federal tax-transfer system is defined as 100% multiplied by one minus the ratio of net household income volatility to gross household income volatility.

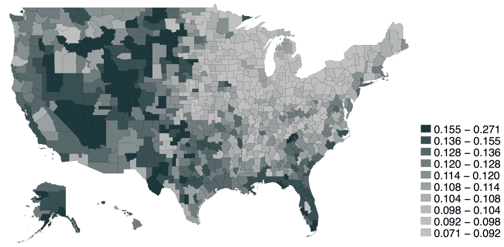
Geographical variation in income volatility

In Figure 2, we present the geographic distribution of the volatility measure when measured separately for each commuting zone in the United States. We see that income volatility tends to be higher on the coasts and in the Western part of the country, but lower in the Midwest and along the Great Lakes. The patterns are broadly similar across income definitions with a correlation of 0.78 between measures of local income volatility in earnings and gross household income and a correlation of 0.99 between between measures of local income volatility in gross and net household income.

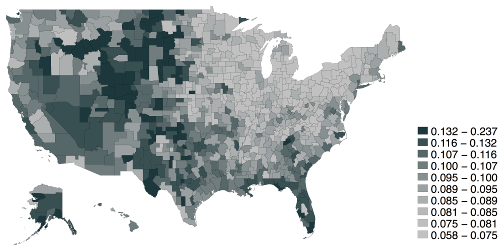
Figure 3(a) presents correlations between gross income volatility and other local socio-economic conditions within the commuting zone, where the measures of commuting zone conditions are from Chetty et al. (2015). Figure 3(b) presents these correlations for net income volatility. Correlations are presented in absolute value, with the sign of the correlation indicated with a symbol of (+) for positive or (-) for negative. Overall, income volatility tends to be larger in areas that are worse on other measures of economic and social conditions. Among economic conditions, income volatility is most positively correlated with local income inequality (as measured by the Gini coefficient) and the local poverty rate. It is negatively related to labor force participation and the fraction of the population with income between the 25th and 75th percentile (a proxy for the middle class). Among social conditions, income volatility is most negatively related to the social capital index (which measures social resource availability) as well as to our measures of short commute time to work, the marriage rate and the college graduation rate. Moreover, local income volatility is positively related to the measures of violent crime rate, segregation experienced by the impoverished, and the high school drop out rate.



(a) Earnings Volatility



(b) Gross Income Volatility



(c) Net Income Volatility

Figure 2: Geographic Distribution of Earnings and Income Volatility

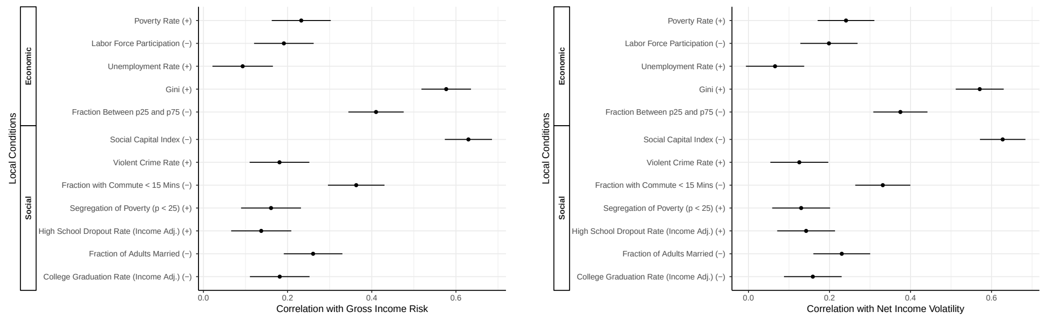
Notes: These figures present the geographic distribution of volatility in individual earnings, gross household income, and net household income. Volatility is defined as movements up or down in a household's income over time, as measured by the variance of the year-by-year changes in log household (annual) income.

4 Causes and consequences of income volatility

In this section, we use several complementary approaches to examine causes and consequences of income volatility.

4.1 Income processes

In this subsection we follow a literature which explores income volatility by estimating a statistical process of income (see e.g. Blundell et al., 2015, Meghir and Pistaferri, 2011). These analyses use the stayers sample. The estimated income process permits a decomposition of the measure of income volatility into various components that capture different sources of volatility. Appendix B.1 lays out and explains the income process and shows how it is identified and estimated. As explained in this appendix, one component of the income process reflects the trends in income over time and across ages that are common to households, whereas the remaining volatility in income can come from at least three sources: idiosyncratic permanent shocks to the worker's income; idiosyncratic transitory shocks to the worker's income; and changes in the firm's performance or productivity (as measured by the value added) that are passed on to workers within that firm.



(a) Gross Income (b) Net Income

Figure 3: Geographic Correlates of Income Volatility

Notes: This figure presents the correlations between (a) local gross household income variability and other local conditions, and (b) local gross household income variability and local conditions. The local conditions are divided into economic and social categories. The data on commuting zone conditions is from Chetty et al. (2015). Commuting zones are weighted by the number of FTE worker observations with valid gross household income measures when estimating the correlations and their standard errors. Correlations are presented in absolute value, with the sign of the correlation indicated with a symbol of (+) for positive or (-) for negative.

Results from baseline model

Appendix B.2 presents the parameter estimates, and reports results from several specification checks. In Table 2 under the column labeled "Firm Only," we summarize the estimation results from the baseline model and use these estimated parameters to quantify the contribution to income volatility from the various sources. The full set of parameter estimates is available in Appendix Table A.2. The standard deviation in log earnings growth is 0.17. Decomposing the variance in log earnings growth, we find that almost 40 percent is due to permanent shocks at the worker level, 58 percent can be attributed to transitory shocks at the worker level, and 3 percent is due to the pass-through of permanent shocks to value added at the firm level .

The estimated pass-through rate $\hat{\gamma}$ is 0.14, suggesting that a 10 percent permanent increase in the value added of the firm leads to a 1.4 percent permanent increase in the earnings of incumbent workers. This pass-through rate is in the range of estimates reported by Card et al. (2018). While permanent shocks to value added are transmitted to workers' earnings, transitory firm shocks are not. This finding is consistent with previous work (see e.g. Guiso et al., 2005; Friedrich et al., 2019). A natural interpretation of this finding is that transitory changes in value added reflect measurement error that does not give rise to economic responses. In the remainder of the report, we will treat the transitory changes in value added as measurement error and focus on the pass-through of the permanent shocks.

We repeat the analysis for gross household income and net household income. Appendix Figure A.3 summarizes the estimation results and uses the estimated parameters to quantify their contributions to volatility in gross and net household income. The standard deviation in log gross income growth is 0.23, while it is 0.21 for log net income growth. We find that 33 percent of gross income volatility is due to permanent shocks at the worker level, 65 percent can be attributed to transitory shocks at the worker level, and 1 percent is due to the pass-through

	Parameters and Growth Decomposition			
	Firm Only		Accounting for Markets	
	Parameter	Var. (%)	Parameter	Var. (%)
Permanent Worker Shock (Std. Dev.)	0.10 (0.00)	39.5%	0.10 (0.00)	38.1%
Transitory Worker Shock (Std. Dev.)	0.13 (0.00)	57.6%	0.13 (0.00)	57.4%
Permanent Firm Shock Passed-through (Std. Dev.)	0.03 (0.00)	2.8%	0.02 (0.00)	1.8%
— Permanent Firm Shock Passthrough Coefficient	0.14 (0.01)		0.13 (0.01)	
Transitory Firm Shock Passed-through (Std. Dev.)	0.00 (0.00)	0.0%	0.00 (0.00)	0.0%
— Transitory Firm Shock Passthrough Coefficient	-0.01 (0.01)		0.00 (0.00)	
Market Shock Passed-through (Std. Dev.)			0.02 (0.00)	1.1%
— Market Shock Passthrough Coefficient			0.18 (0.02)	

Table 2: Variance Decomposition in Baseline Passthrough Specification

Notes: This table presents the variance decomposition of the joint process for growth in value added and earnings in the baseline specification. In the estimation sample, the standard deviation of log value added growth is 0.31, and the standard deviation of log earnings growth is 0.17. The estimated moving average coefficients are used to construct the variances of transitory shocks. The full set of parameter estimates is available in Panel A of Appendix Table A.2.

of permanent shocks to value added at the firm level; these percentages are very similar for log net household income. We also find pass-through rates of firm permanent shocks to gross and net household income of 0.136 and 0.127, respectively. This indicates that the progressive nature of the Federal tax-transfer attenuates more than 10 percent of the economic consequences of a firm shock.

Contribution of firms versus markets

We have thus far followed the existing literature in assuming that all income volatility is specific to the individual or the household, abstracting from aggregate regional or industry shocks. We now take two steps to distinguish between shocks that are specific to individual workers versus those that are common to workers in a market.

First, we estimate the earnings and value added processes conditional on a full set of year times market fixed effects. The results are presented in Table 2 under the column labeled "Net of Market." Decomposing the variance in log earnings growth within markets, we find that 38 percent is due to permanent shocks at the worker level, 57 percent can be attributed to transitory shocks at the worker level, and 2 percent is due to the pass-through of permanent shocks to value added at the firm level. Conditional on the full set of year times market fixed effects, the estimated pass-through rate $\hat{\gamma}$ is 0.13. By comparison, the estimated pass-through rate of permanent market shocks is as large as 0.18. This finding highlights the importance of distinguishing between shocks that are specific to workers in a given firm versus those that common to workers in a market. In Appendix Figure A.3, we control for a full set of year times market fixed effects when analyzing gross household income and net household income, finding also here an important role for shocks that are common to the market.

The second approach we use to examine the variability across regions and industries in income volatility (and its sources) is to estimate the earnings and value added processes separately for each broad market. This estimation generates market-specific parameters for the processes. We perform this analysis separately for earnings, gross income, and net income. The results are presented in Appendix Figure A.4. These results reveal that a pass-through rate of log value added permanent shocks to workers' earnings that is higher in the goods sector as compared to the services sector. However, the Federal tax-and-transfer system attenuates some of these differences.

Robustness checks

Appendix Figure A.2a explores how the pass-through rate varies across worker types by estimating the earnings and value added processes separately for each subgroup. Conditional on a full set of year times market fixed effects, we find that the pass-through estimate does not vary much by the worker's age, previous wage, or gender. Moreover, the pass through rates do not change materially if we restrict the sample to workers who were first hired at the firm in the beginning of the eight year employment spell versus those that have stayed in the firm for a longer time.

In Appendix Figure A.2b, we also present results from several other specification checks. Following Guiso et al. (2005), our main measure of firm performance is value added. They offer two reasons for using value added as a measure of firm performance. First, they argue, value added is the variable that is directly subject to stochastic fluctuations. Second, firms have discretionary power over the reporting of profits in balance sheets, which makes profits a less reliable objective to assess. Nevertheless, it is reassuring to find that the estimates of the pass-through rates are broadly similar if we measure firm performance by operating profits, earnings before interest, tax and depreciation (EBITD), or value added net of reported depreciation of capital. We also show that the estimated pass-through is in the same range as our baseline result if we exclude multinational corporations (for which it can be difficult to accurately measure value added) or exclude the largest firms (which are more likely to have multiple plants).

Our analyses so far have relied on statistical processes of earnings and value added. While this is common, the identification of shocks and pass-through rates relies on plausible but ultimately debatable identifying assumptions. One concern is that individuals may change their behavior in important ways in response to income shocks. For example, an exogenous increase in income could cause workers to decrease labor supply, which could lead us to underestimate the pass-through to workers' earnings of firm shocks. More indirect sources of bias in the estimation of pass-through to workers could arise if individuals respond on margins that are correlated with labor supply and earnings, such as capital investment, expenditure, marriage and fertility decisions, geographic mobility, health investments, and retirement. To assess this, we analyze information in Form W-2G on state lottery winnings. In particular, we leverage variation in the timing of a lottery win, and form cohorts of lottery winners that win in different years. We observe that prior to winning the lottery, the trend in employment, earnings, and other outcomes evolve very similarly across lottery winning cohorts. This suggests a quasi-experimental research design where we use later winners in the years prior to their lottery win as a control group for current winners in the same calendar years. Concretely, this research design amounts to a difference-in-differences analysis, where we use future winners to net out changes in the outcome of interest due to common macroeconomic/time effects as well as common effects of the passage of time, with the underlying identifying assumption that the exact timing of the lottery win is unrelated to pre-existing outcome trends. We develop this identification strategy in greater detail in Appendix D.1.

In Appendix D.1, we summarize difference-in-differences estimates of the the behavioral responses to exogenous changes in income induced by lottery winnings in terms of earnings, employment, and related behavioral response margins. Reassuringly, we find modest effects on both employment and earnings per dollar won. On average, prize winners reduce their employment by roughly \$0.02 per dollar won. After estimating the average effects of income changes induced by lottery winnings, we explore heterogeneity in impacts in various dimensions such as age and prize size. Furthermore, Appendix D.4 summarizes related results, but in terms of per-period income, where we use two distinct approaches to allocating one-time lottery income shocks into annual income shocks.

Taken together, our analysis of lottery-induced changes in income lends support to the modeling of the income process. One remaining concern, however, is that the identification of firm shocks and pass-through rates rely on plausible but ultimately debatable identifying assumptions. To critically examine and relax these assumptions – and thereby improve the quality and credibility of our analyses – we extract observable firm shocks based on outcomes of public procurement auctions run by state governments, particularly departments of transportation (DOTs). When two firms bid blindly for the same contract, and one firm happens to win by a small margin, a (quasi)experiment is produced. The winning firm receives an as good as random change in the demand for their products or services, producing exogenous shocks to their labor demand. Using many such experiments from state government DOT public auction data, we examine how these shocks transmit to workers.

We use the experiment of winning a procurement auction to estimate the pass-through of firm-specific shocks. To do so, we match the records of firms that bid in procurement auctions to their tax information using a matching algorithm. Appendix Table A.4(a) demonstrates that the matching algorithm performs well in validation exercises. Appendix Table A.4(b) provides an overview of the sample and shows that it is representative of the broader economy. Nationally, the firms that were matched to procurement auctions represent around 10% of all EBITD and employment in the construction industry. Appendix Table A.4(c) provides basic sample characteristics on the main outcome variables of interest.

Given this data, Appendix Figures A.5(a-b) provide the pass-through effects of winning a procurement auction on log earnings per worker and log number of employees, respectively. We estimate these effects both prior to the auction announcement ("Before") and after the auction winner is announced ("After"). Since the auction should not affect the firm's demand for labor prior to the auction winner being announced, we expect to find no effects during the Before period, so the Before period serves as a falsification test. In the bars labeled "Baseline", we compare auction winners to non-recipients that had never won an auction before and placed a bid at the same time as the winners but lost. In the After period, we find that earnings per worker increase by about 2% while the number of employees increases by about 8%. The ratio of the effect on log employment to the effect on log earnings per worker recovers the labor supply elasticity, which is presented in the "Baseline" bar of Appendix Figure A.8 to be about 4. We show in Appendix Figure A.5 that these effects are similar if inferring the labor supply elasticity from log receipts and using the shift-share design. We then use this labor supply elasticity estimate to fit the rent-sharing model of Kroft et al. (2021). In Appendix Table A.41, we find similar estimates when using GMM or OLS estimation for the key parameters, which are the product demand elasticity $1/\epsilon$, the composite returns to labor ρ, the marginal returns to labor β_L, and the interquartile range of TFP estimates. Appendix Figure A.30 demonstrates that these parameters are relatively similar across broad markets and years.

For robustness of the estimates based on procurement auctions, we consider a number of alternative specifications. In Appendix Figures A.5(a-b), we consider restricting the control sample to firms that had never bid in an auction before (bar labeled "Sample: First-timers"), firms that would not go on to win an auction in the future (bar labeled "Sample: Never-winners"), or firms that bid in the same auction as the winners (bar labeled "Sample: Same Auction"), finding very similar effects in the After period and passing the falsification test in the Before period. For the effect on mean log earnings in Appendix Figures A.5(a), we also consider re-estimating these effects on the mean earnings of stayers or workers with longer tenure in the firm, finding similar effects. The corresponding bars in Appendix Figure A.8 demonstrate that the labor supply elasticity is not sensitive to these alternative specifications. Appendix Figure A.6 demonstrates that the results are relatively stable across years after the auction occurs. Appendix Figure A.7 demonstrates that the results are similar regardless of whether or not the state in which the auction occurs has right-to-work or prevailing wage laws. Appendix Figure A.9(a-b) demonstrates that the robustness of the results is not sensitive to the number of years used when defining stayers and tenured workers, respectively. Appendix Figure A.9(c) demonstrates that the effects on stayers are not sensitive to the minimum earnings threshold required for workers to be included in the sample. Appendix Figure A.9(d) demonstrates that, when we restrict the sample to winners and losers whose bids were within a close bandwidth of one another, the results remain similar.

4.2 Mover analyses

So far, we focused on income volatility among workers who stay in the same firm over time. We now turn attention to an approach which examines income changes associated with a worker entering a new firm.

In the baseline results, we consider a special case which assumes that $\phi_{ij} = x_i + \psi_j$ and that $\gamma = \Upsilon = 0$. The first restriction imposes a log additive structure on the earnings that worker i can expect to receive from working in firm j. Under this functional form, the worker fixed effect captures the (time-invariant) portable component of earnings ability, whereas the firm fixed effect can be interpreted as a firm-specific relative pay premium. The second restriction assumes there is no pass through of firm or market level shocks. As a result, the firm effects on earnings do not vary over time. By invoking these two restrictions, our statistical model of earnings reduces to the two-way (worker and firm) fixed effect model of AKM. Appendix C.2 presents results from relaxing these assumptions, and Appendix C.3 provides additional robustness checks.

Under the above restrictions, the variance of log earnings can be written as:

$$Var(\log W_{it}) = \underbrace{Var(x_i + \mathcal{X}'_{it}b)}_{\text{Worker component}} + \underbrace{Var(\psi_{j(i,t)})}_{\text{Firm component}} + \underbrace{2Cov(x_i + \mathcal{X}'_{it}b, \psi_{j(i,t)})}_{\text{Sorting component}} + \underbrace{Var(\epsilon_{it})}_{\text{Residual}} \quad (1)$$

where the worker and firm components tell us how much of the variation in log earnings can be attributed to heterogeneity in worker and firm effects, respectively. The third component captures the contribution to earnings inequality from the sorting of workers to firms. The goal is to quantify these three components to draw inference about the determinants of earnings inequality in the U.S. economy. The decomposition includes both workers who move between firms and stayers. However, the firm and worker effects are only separately identified within a connected set of firms that are linked by worker mobility. Consistent with previous work, we therefore restrict our sample of workers (including stayers and movers) to those who work at a firm in the largest connected set in each time interval (2001-2008 and 2008-2015). In the U.S., this set covers more than 90 percent of the workers (see Appendix Table A.6).

In Table 3, we present results from the variance decomposition in (1) based on data for all firms and workers in the connected set (which includes both workers who move between firms and stayers). This table reports estimates of the worker, firm and sorting components as defined in equation (1). Appendix Table A.7 shows estimates of the subcomponents in the second equality of (1). Consider first Panel A of Table 3 where we present estimates from the AKM estimator for two different time periods (2001-2008 and 2008-2015) as well as pooled estimates where we combine the data from these time periods. The results show that the worker, firm and sorting components change little over time. Therefore, we focus attention on the pooled estimates. These results suggest that the firm effects explain around 9 percent of the variation in log earning, whereas worker sorting accounts for 5 percent. The correlation between firm effects and worker effects is only 0.1.

Next, consider Panel B of Table 3 where we report the BLM estimates. As discussed in Appendix C.1, a possible advantage of the BLM estimator is that it addresses limited mobility bias. Once we correct for such bias we find that firm effects are very small in the U.S. labor market, accounting for only 3 percent of the variation in log earnings. Instead, a larger part of the earnings variation is explained by worker sorting. The correlation between firm effects and

Years:		2001-2008	2008-2015	Pooled
Panel A.		**AKM Estimation**		
Share explained by:				
i) Worker Effects	$Var(x_i)$	75%	75%	75%
ii) Firm Effects	$Var(\psi_{j(i)})$	9%	9%	9%
iii) Sorting	$2Cov(x_i, \psi_{j(i)})$	5%	6%	5%
Sorting Correlation:	$Cor(x_i, \psi_{j(i)})$	0.09	0.11	0.10
Panel B.		**BLM Estimation**		
Share explained by:				
i) Worker Effects	$Var(x_i)$	72%	72%	72%
ii) Firm Effects	$Var(\psi_{j(i)})$	3%	3%	3%
iii) Sorting	$2Cov(x_i, \psi_{j(i)})$	13%	14%	14%
Sorting Correlation:	$Cor(x_i, \psi_{j(i)})$	0.43	0.46	0.44

Table 3: AKM and BLM Log Earnings Decomposition Estimates

Notes: This table presents the decomposition of log earnings variation using the AKM and BLM estimators for two time periods.

worker effects exceeds 0.4 once we correct for limited mobility bias. This finding suggests that sorting of better workers to better firms is an empirically important feature of the U.S. labor market. Detailed sorting patterns are presented in Appendix Figure A.13.

In Table 4, we repeat the AKM and BLM analyses to understand the roles of firm effects, worker effects, and the sorting of workers to firms in explaining gross household income and net household income. We find that moving to a new firm causes even smaller changes in gross and net income as compared to earnings. Moreover, sorting of better workers to better firms contribute less to inequality in gross and net income than to dispersion of earnings. By contrast, worker effects explain a larger part of the variation in gross and net income as compared to gross earnings. Finally, the sorting of workers to firms explains about 5% of the variance in each income measure compared to 9% for earnings, indicating that the Federal tax-and-transfer system attenuates the incentives for better workers to move to better firms.

Inequality within and between firms

We now shift attention to describing the inequality within and between firms. To do so, we follow Song et al. (2018) in expressing the variance of log earnings as:

Income Measure:		Earnings	Gross Income	Net Income
		Panel A. AKM Estimation		
Share explained by:				
i) Worker Effects	$Var(x_i)$	75.4%	83.6%	84.4%
ii) Firm Effects	$Var(\psi_{j(i)})$	8.8%	5.0%	4.7%
iii) Sorting	$2Cov(x_i, \psi_{j(i)})$	4.9%	2.6%	2.2%
Sorting Correlation:	$Cor(x_i, \psi_{j(i)})$	0.09	0.06	0.05
		Panel B. BLM Estimation		
Share explained by:				
i) Worker Effects	$Var(x_i)$	72.4%	81.0%	84.0%
ii) Firm Effects	$Var(\psi_{j(i)})$	3.2%	0.4%	0.4%
iii) Sorting	$2Cov(x_i, \psi_{j(i)})$	12.9%	5.4%	4.8%
Sorting Correlation:	$Cor(x_i, \psi_{j(i)})$	0.43	0.45	0.44

Table 4: AKM vs BLM by Income Measure

Notes: This table presents AKM and BLM decomposition estimates for log earnings, gross income, and net income.

$$\begin{aligned}
Var(\log W_{it}) &= \underbrace{Var\left(\log W_{it} - \mathbb{E}\left[\log W_{it} | j(i,t) = j\right]\right)}_{\text{Within}-\text{firm}} + \underbrace{Var\left(\mathbb{E}\left[\log W_{it} | j(i,t) = j\right]\right)}_{\text{Between}-\text{firm}} \qquad (2) \\
&= \underbrace{Var\left(x_i + \mathcal{X}_{it}'b - \mathbb{E}\left[x_i + \mathcal{X}_{it}'b | j(i,t) = j\right]\right)}_{\text{Worker heterogeneity within firms}} + \underbrace{Var(\epsilon_{it})}_{\text{Residual}}, \\
&+ \underbrace{Var\left(\psi_{j(i,t)}\right)}_{\text{Firm effects}} + \underbrace{2Cov\left(x_i + \mathcal{X}_{it}'b, \psi_{j(i,t)}\right)}_{\text{Sorting}} + \underbrace{Var\left(\mathbb{E}\left[x_i + \mathcal{X}_{it}'b | j(i,t) = j\right]\right)}_{\text{Segregation}}
\end{aligned}$$

where the first equality expresses the variance of log earnings in terms of inequality within and between firms, and the second equality decomposes these terms into economically interpretable subcomponents. Our interest is centered on the last three subcomponents, which capture distinct sources of inequality between firms: dispersion of firm pay premiums ("Firm effects"); sorting of high earning workers into high paying firms ("Sorting"); and worker segregation which reflects differences in the quality of the workforce across firms ("Segregation"). Both worker sorting and segregation reflect non-random allocation of workers to firms. However, sorting matters for aggregate inequality, whereas segregation does not. This is because an increase in segregation will be offset by a reduction in within-firm inequality. Thus, changes in segregation by itself does not affect earnings inequality; it does, however, matter for the relative importance of inequality within versus between firms.

To perform the decomposition in (2), we use exactly the same sample as in Table 3 which includes both workers who move between firms and stayers. The results are presented in Table 5. In Panel A, we report the terms in the first equality. We find that around one-third of the variance of log earnings can be accounted for by the dispersion of average earnings between firms. The remainder is due to heterogeneity across workers within firms. A comparison of the estimates across the two first columns suggests the between firm component has become slightly more important for inequality over time. This finding is broadly consistent with the results reported in Song et al. (2018).[4]

In the next two panels of Table 5, we use the procedures of AKM and BLM to estimate the subcomponents from the second equality. There are three main findings from this analysis. First, a vast majority of the inequality within firms can be accounted for by the observable characteristics and the fixed effects of the workers. Indeed, only 16 percent of the within-firm inequality reflects time-varying unobservables of the worker. Second, once one addresses limited mobility bias then firm effects explain only 10 percent of the inequality between firms. By comparison, sorting of high earning workers to high paying firms accounts for 40 percent while the remaining 50 percent can be attributed to worker segregation that is unrelated to firm pay premiums. Third, there seems to be little if any changes in the relative importance of firm effects, worker sorting and segregation over the time intervals we consider.

Our finding of the inequality contribution from firm effects changing little over time is consistent with Song et al. (2018), albeit their analysis uses AKM and thus suffers from limited mobility bias. Table 5 reveals, however, that this bias does not change materially over the time intervals we consider. As a result, bias correction seems to be empirically important for accurately describing the cross-sectional distribution of earnings in the U.S., but not for understanding the growth in earnings inequality.[5]

Analyses of firm entry and exit

While the analyses discussed above allow us to understand the income volatility due to shocks to or mobility between existing firms, they do not tell us how workers are affected by entry or exit of firms in the same location. Intuitively, when a large factory opens or shuts down, we expect it to impact the commuting zone as a whole rather than only its own workers, affecting important economic outcomes like income variability, tax payments, and the unemployment rate. As explained in Appendix E.1, we obtain an instrumental variable for firm entry and exit with plausibly exogenous variation by analyzing how aggregate fluctuations in foreign economies may affect foreign-owned firms' decisions to enter or exit a location, making use of the information on foreign ownership. In particular, when a foreign economy expands or contracts, we expect it to result in more entry or exit of foreign-owned firms in the commuting zones with higher initial concentration of activity by owners from that country, allowing us to draw causal inference (under plausible identifying assumptions) by comparing those locations that do and do not receive these shocks.

[4]The analyses in Song et al. (2018) is based on data from 1978 to 2013. Over this longer time period, they show that earnings inequaliy increased considerably, primarly due to a significant rise in the dispersion of average earnings across firms. During the period we consider, however, Song et al. (2018) also report a modest increase in earnings inequality, overall and between firms.

[5]Song et al. (2018) also argue that increases in sorting and segregation caused a large increase in between-firm inequality from 1981 to 2013. At first sight, it would seem like this is inconsistent with our findings. However, most of these increases happen before our data start. During the intervals since 2001 that we consider, Song et al. (2018) report modest increases in the contributions to between-firm inequality from sorting and segregation and a modest decrease from firm effects, consistent with our AKM estimates.

Years:		2001-2008	2008-2015	Pooled
Panel A.		**Total Decomposition**		
Within Firm Share:	$Var(w_{it} - \mathbb{E}[w_{it}\|j])$	67%	64%	66%
Between Firm Share:	$Var(\mathbb{E}[w_{it}\|j])$	33%	36%	34%
Panel B.		**AKM Decomposition**		
Shares of Within Firm Variance:				
Worker Heterogeneity:	$Var(x_i + X'_{it}b - \mathbb{E}[x_i + X'_{it}b\|j])$	84%	85%	84%
Residual:	$Var(\epsilon_{it})$	16%	15%	16%
Shares of Between Firm Variance:				
Firm Effects:	$Var(\psi_j)$	27%	25%	26%
Segregation:	$Var(\mathbb{E}[x_i + X'_{it}b\|j])$	58%	59%	59%
Sorting:	$2Cov(x_i + X'_{it}b, \psi_j)$	15%	16%	15%
Panel C.		**BLM Decomposition**		
Shares of Within Firm Variance:				
Worker Heterogeneity:	$Var(x_i + X'_{it}b - \mathbb{E}[x_i + X'_{it}b\|j])$	83%	84%	84%
Residual:	$Var(\epsilon_{it})$	17%	16%	16%
Shares of Between Firm Variance:				
Firm Effects:	$Var(\psi_j)$	10%	10%	10%
Segregation:	$Var(\mathbb{E}[x_i + X'_{it}b\|j])$	50%	50%	50%
Sorting:	$2Cov(x_i + X'_{it}b, \psi_j)$	40%	40%	40%

Table 5: AKM and BLM Within and Between Inequality Decompositions

Notes: This table presents the decomposition of log earnings variation within and between firms using the AKM and BLM estimators for two time intervals. The analysis uses both workers who move between firms and stayers.

In Appendix E.2, we present and discuss the parameter estimates and the results from a number of specifications and robustness checks. As shown in Appendix Table A.26, we find a positive and statistically significant effect of a firm entering the market on the earnings of workers at existing firms in the same commuting zone. To put the estimates in context, we find that, if a firm employing 10 percent of employees in the commuting zone exits and lays off its workers, then workers in existing employment relationships at other firms in the same commuting zone experience a 4.5% decline in employment, a 4.7% decline in earnings payments to workers, and a 6.4% decline in the firm's value added. Interestingly, as shown in Appendix Table A.32, we find that a layoff shock due to firm exit results in a substantial decrease in gross and net income among workers at other firms in the same commuting zone. However, the impact on net income is smaller than the effect on earnings and gross income, and the estimates imply that the Federal tax-and-transfer system provides about a 9 percent rate of insurance against shocks due to other firms entering and exiting.

4.3 Income volatility, tax revenues, and receipt of tax credits

So far, we have focused on documenting income volatility, quantifying its sources, and examining the attenuation from the Federal tax-and-transfer system. We now shift attention to examining how income volatility, at the individual and market level, may make it difficult to predict tax revenues or receipt of tax credits.

Income processes and mover analyses

The goal is to use the estimates from the income processes and mover analyses to i) examine how various sources of income volatility may generate change in tax revenues, and ii) illustrate how the impact on tax revenues of income volatility may depend on the progressivity of the tax system. As a first step, however, it is convenient to parametrize the tax schedule. Following Heathcote et al. (2014) and Blundell et al. (2016), we choose the following log-linear parametrization to approximate the effective tax rates implicit in the Federal tax-and-transfer system.

$$\tilde{I}_{i,t} = \tau I_{it}^{\lambda}$$

where I denotes gross income and $\tilde{I}$ denotes net income. We estimate these parameters outside the model. In each year, we regress log net household income (earnings plus other income minus taxes) on log household gross income (earnings plus other income) for our sample. The construction of these income measures is detailed in Appendix A. The intercept from this regression gives us τ while λ is identified from the slope coefficient. We estimate τ of around 0.89 whereas λ is estimated to be about 0.92.[6] In a proportional tax-transfer system, λ is equal to one and $(1-\tau)$ is the proportional effective tax rate. By contrast, if $0 < \lambda < 1$, then the marginal effective tax rate is increasing in earnings. Appendix Figure A.17 shows how well our parsimonious tax function approximates the effective tax rates implicit in the complex U.S. tax-transfer system. Here we compare the predicted log net income from the regression to the observed log net income across the distribution of log gross income, finding that this specification provides an excellent fit.

First, we use the estimated tax schedule to understand how firm shocks are passed-through to tax revenues. To do so, we simulate a one standard deviation shock to log value added at the firm. Then, we use the estimated passthrough rates and the estimated tax-transfer system to collect the implied changes in gross and net income, which in turn provide us the average tax revenue response to a firm shock. Mean tax revenues rise in response to a firm shock in the baseline tax system, as all workers have greater income and marginal tax rates are positive. Finally, we change the parameter λ in order to investigate how the average tax revenue response to the firm shocks depends on tax progressivity. Figure 4 presents the results of this exercise. We find that the responsiveness of tax revenues to firm shocks is greater when the tax schedule is more progressive.

Second, we use the estimated tax schedule to understand how sorting across firms affects mean tax revenues. To do so, we use the AKM model for log gross income (it is the sum of the firm effect, the worker effect, and the worker-year residual) but randomly re-assign firm effects to construct log gross income without sorting. We then use the tax function to collect

[6] These results mirror closely existing U.S. estimates of τ and λ (see e.g. Guner et al., 2014, Heathcote et al., 2017).

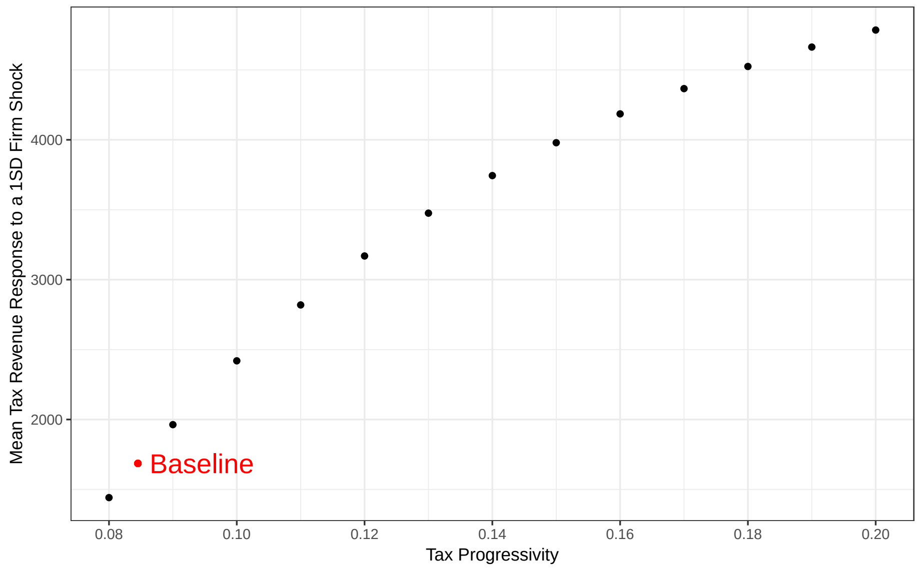


Figure 4: Mean Tax Revenue Responses to a Passed-through Firm Shock, by Tax Progressivity

Notes: In this figure, we consider the effect of a firm shock on mean tax revenues. To do so, we simulate a one standard deviation shock to log value added at the firm. Then, we use the estimated passthrough rates and the estimated tax-transfer system to collect the implied changes in gross and net income, which in turn provide us the average tax revenue response to a firm shock. Finally, we change the parameter λ in order to investigate how the average tax revenue response to the firm shock depends on tax progressivity.

implied net income and tax revenues for each worker, with and without sorting. Mean tax revenues fall without sorting in the baseline tax system, as fewer workers receive high incomes and high incomes face greater marginal tax rates. Finally, we change the parameter λ in order to investigate how the average tax revenue response to sorting depends on tax progressivity. Figure 5 presents the results of this exercise. We find that the tax revenue gains from sorting are greater when the tax schedule is more progressive.

Analyses of firm entry and exit

Appendix Table A.32 uses the instrumental variables design described in Appendix E.1 in order to estimate the effect that firm entries and exits has on tax payments made by workers employed at other firms in the same commuting zone. It finds a positive and statistically significant effect. To put the estimates in context, we find that, if a firm employing 10 percent of employees in the commuting zone exits and lays off its workers, then workers in existing employment relationships at other firms in the same commuting zone experience a 6.8 percent decline in tax payments.

In order to better understand the responsiveness of tax revenues, we highlight an important component of the system, the Earned Income Tax Credit (EITC). We consider two margins of EITC utilization – claiming any EITC deduction (the extensive margin) and the amount of deduction claimed (the intensive margin).[7] In Appendix Table A.32, we find negative effects of

[7] Because the EITC is zero either at t or $t-1$ for an observation that experiences a change in the EITC extensive margin, we cannot explore log differences, as we do for other outcomes. Instead, we consider the

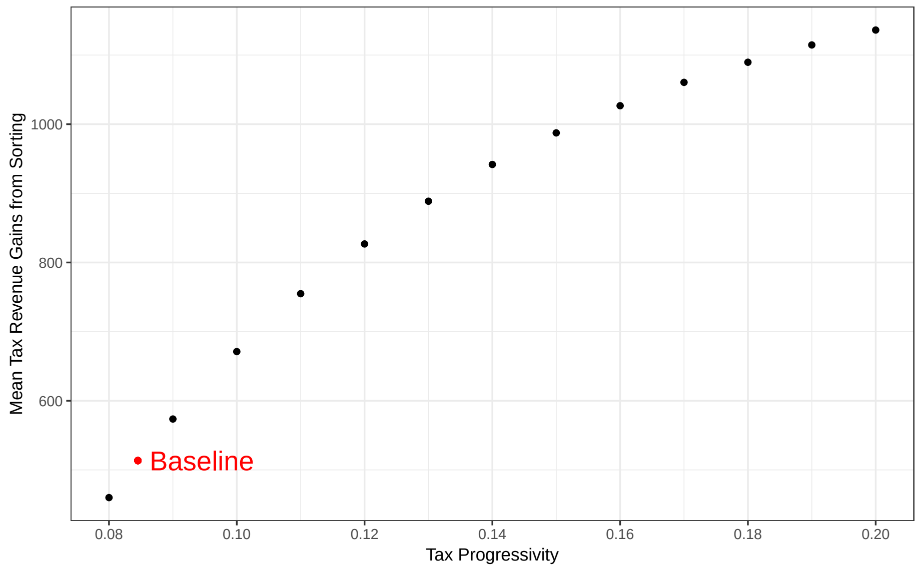


Figure 5: Mean Tax Revenue Gains from Sorting, by Tax Progressivity

Notes: In this figure, we consider the gains from sorting on mean tax revenues. To do so, we randomly re-assign firm effects across workers so that firm effects and worker effects are uncorrelated, then reconstruct log gross income. Then, we use the estimated passthrough rates and the estimated tax-transfer system to collect the implied changes in gross and net income, which in turn provide us the average tax revenue gains from sorting. Finally, we change the parameter λ in order to investigate how the average tax revenue gains from sorting depends on tax progressivity.

firm entry on both the extensive and intensive margin of EITC participation, though statistical precision is somewhat limited. To put the estimates in context, we find that, if a firm employing 10 percent of employees in the commuting zone exits and lays off its workers, then workers in existing employment relationships at other firms in the same commuting zone experience approximately a 2.4 percent increase in EITC take-up and approximately a 3.1 percent increase in the EITC deduction claimed. These results suggest that the EITC is an active channel through which tax revenues adjust to insure against firm entry and exit shocks for workers employed at other firms in the commuting zone.

5 Concluding remarks

In this report, we documented income volatility and examined its causes and consequences. Our empirical findings raise questions such as: What do small firm effects, strong sorting and significant pass-through of firm shocks tell us about the functioning of the labor market? How would changes in tax policy affect earnings inequality, worker sorting, income volatility and tax payments? To answer these questions, we have developed an equilibrium model of the labor market that can match the empirical findings that we documented in this report. The model not

transformation of Davis et al. (1996), which is able to include any observation that experiences an extensive margin change, and can be interpreted as an approximation to the log difference for small changes; see their paper for further details.

only allows us to economically interpret the empirical findings reported here but also to improve the analysis of income volatility and taxation. In particular, our model captures that changes in tax policy may induce firms to change their hiring and wage setting, and such changes may affect workers' choice of firm, industry and region. Thus, we can perform model-based simulations of tax reforms which relax the assumption made in Section 4.3 of no behavioral responses of workers and firms to the changes in tax progressivity. See Appendix F for additional results based on the equilibrium model of the labor market.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008): "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?" *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 171, 673–697.

Balke, N. and T. Lamadon (2020): "Productivity Shocks, Long-Term Contracts and Earnings Dynamics," *Working paper.*

Barth, E., A. Bryson, J. C. Davis, and R. Freeman (2016): "It's where you work: Increases in earnings dispersion across establishments and individuals in the US," *Journal of Labor Economics*, 34, S67–S97.

Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016): "Female Labor Supply, Human Capital, and Welfare Reform," *Econometrica*, 84, 1705–1753.

Blundell, R., M. Graber, and M. Mogstad (2015): "Labor income dynamics and the insurance from taxes, transfers, and the family," *Journal of Public Economics*, 127, 58–73.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87, 699–739.

Borovickova, K. and R. Shimer (2017): "High Wage Workers Work for High Wage Firms," Working Paper 24074, National Bureau of Economic Research.

Burchardi, K., T. Chaney, and T. Hassan (2016): "The effect of migration on foreign direct investment," *Working paper.*

Card, D. (2001): "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 19, 22–64.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 36, S13–S70.

Card, D., F. Devicienti, and A. Maida (2013a): "Rent-sharing, holdup, and wages: Evidence from matched panel data," *The Review of Economic Studies*, 81, 84–111.

Card, D., J. Heining, and P. Kline (2013b): "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly Journal of Economics*, 128, 967–1015.

Carlsson, M., J. Messina, and O. N. Skans (2016): "Wage adjustment and productivity shocks," *The Economic Journal*, 126, 1739–1773.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011): "How does your kindergarten classroom affect your earnings? Evidence from Project STAR," *The Quarterly Journal of Economics*, 126, 1593–1660.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2015): "The Economic Impact of Tax Expenditures: Evidence from Spatial Variation Across the U.S." *SOI Working Paper*.

Davis, S. J., J. Haltiwanger, and S. Schuh (1996): "Small business and job creation: Dissecting the myth and reassessing the facts," *Small business economics*, 8, 297–315.

DeBacker, J., B. Heim, V. Panousi, S. Ramnath, and I. Vidangos (2013): "Rising inequality: transitory or persistent? New evidence from a panel of US tax returns," *Brookings Papers on Economic Activity*, 2013, 67–142.

Dunne, T., L. Foster, J. Haltiwanger, and K. R. Troske (2004): "Wage and productivity dispersion in United States manufacturing: The role of computer investment," *Journal of Labor Economics*, 22, 397–429.

Eeckhout, J. and P. Kircher (2011): "Identifying Sorting In Theory," *The Review of Economic Studies*, 78, 872–906.

Faggio, G., K. G. Salvanes, and J. Van Reenen (2010): "The evolution of inequality in productivity and wages: panel data evidence," *Industrial and Corporate Change*, 19, 1919–1951.

Friedrich, B., L. Laun, C. Meghir, and L. Pistaferri (2019): "Earnings Dynamics and Firm-Level Shocks," Working Paper 25786, National Bureau of Economic Research.

Gibbons, R. and L. Katz (1992): "Does unmeasured ability explain inter-industry wage differentials?" *The Review of Economic Studies*, 59, 515–535.

Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2005): "Comparative advantage, learning, and sectoral wage determination," *Journal of Labor Economics*, 23, 681–724.

Guiso, L., L. Pistaferri, and F. Schivardi (2005): "Insurance within the firm," *Journal of Political Economy*, 113, 1054–1087.

Guner, N., R. Kaygusuz, and G. Ventura (2014): "Income taxation of US households: Facts and parametric estimates," *Review of Economic Dynamics*, 17, 559–581.

Heathcote, J., K. Storesletten, and G. L. Violante (2014): "Consumption and labor supply with partial insurance: An analytical framework," *American Economic Review*, 104, 2075–2126.

——— (2017): "Optimal tax progressivity: An analytical framework," *The Quarterly Journal of Economics*, 132, 1693–1754.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018): "Shift-share instruments and the impact of immigration," *Working paper*.

Kline, P., N. Petkova, H. Williams, and O. Zidar (2019): "Who profits from patents? rent-sharing at innovative firms," *The Quarterly Journal of Economics*, 134, 1343–1404.

Kline, P., R. Saggio, and M. Sølvsten (2020): "Leave-out estimation of variance components," *Econometrica*, 88, 1859–1898.

Kroft, K., Y. Luo, M. Mogstad, and B. Setzler (2021): "Imperfect Competition and Rents in Labor and Product Markets: The Case of the Construction Industry," Working Paper w27325, National Bureau of Economic Research.

Meghir, C. and L. Pistaferri (2011): "Earnings, consumption and life cycle choices," in *Handbook of labor economics*, Elsevier, vol. 4, 773–854.

Murphy, K. M. and R. H. Topel (1990): "Efficiency wages reconsidered: Theory and evidence," in *Advances in the Theory and Measurement of Unemployment*, Springer, 204–240.

Piketty, T. and E. Saez (2003): "Income inequality in the United States, 1913–1998," *The Quarterly Journal of Economics*, 118, 1–41.

Saez, E. and G. Zucman (2016): "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data," *The Quarterly Journal of Economics*, 131, 519–578.

Shimer, R. and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–369.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. von Wachter (2018): "Firming Up Inequality," *The Quarterly Journal of Economics*, 134, 1–50.

Sorkin, I. (2018): "Ranking Firms using Revealed Preference," *The Quarterly Journal of Economics*, 133, 1331–1393.

Yagan, D. (2015): "Capital tax reform and the real economy: The effects of the 2003 dividend tax cut," *American Economic Review*, 105, 3531–3563.

A Appendix: Sample Construction and Variable Definitions

All firm-level variables are constructed from annual business tax returns over the years 2001-2015: C-Corporations (Form 1120), S-Corporations (Form 1120-S), and Partnerships (Form 1065). Worker-level variables are constructed from annual tax returns over the years 2001-2015: Direct employees (Form W-2), independent contractors (Form 1099), and household income and taxation (Form 1040).

Variable Definitions:

- **Earnings:** Reported on W-2 box 1 for each Taxpayer Identification Number (TIN). Each TIN is de-identified in our data.

- **Gross Household Income:** Using a definition similar to that of Piketty and Saez (2003), we define gross household income as the sum of taxable wages and other income (line 22 on Form 1040) minus unemployment benefits (line 19 on Form 1040) minus taxable Social Security benefits (line 20a on Form 1040) plus tax-exempt interest income (line 8b on Form 1040). We at times also consider this measure when subtracting off Schedule D capital gains (line 13 on Form 1040).

- **Federal Taxes on Household Income:** This is given by the sum of two components. The first component is the sum of FICA Social Security taxes (given by 0.0620 times the minimum of the Social Security taxable earnings threshold, which varies by year, and taxable FICA earnings, which are reported on Box 3 of Form W-2) and FICA Medicare taxes (given by 0.0145 times Medicare earnings, which are reported on Box 5 of Form W-2). The second component is the sum of the amount of taxes owed (the difference between line 63 and line 74 on Form 1040, which is negative to indicate a refund) and the taxes already paid or withheld (the sum of lines 64, 65, 70, and 71 on Form 1040).

- **Net Household Income:** We construct a measure of net household income as Gross Household Income minus Federal Taxes on Household Income plus two types of benefits: unemployment benefits (line 19 of Form 1040) and Social Security benefits (line 20a of Form 1040).

- **Employer:** The Employer Identification Number (EIN) reported on W-2 for a given TIN. Each EIN is de-identified in our data.

- **Wage Bill:** Sum of Earnings for a given EIN plus the sum of 1099-MISC, box 7 nonemployee compensation for a given EIN in year t.

- **Size:** Number of FTE workers matched to an EIN in year t.

- **NAICS Code:** The NAICS code is reported on line 21 on Schedule K of Form 1120 for C-corporations, line 2a Schedule B of Form 1120S for S-corporations, and Box A of form 1065 for partnerships. We consider the first three digits to be the industry. We code invalid industries as missing.

- **Commuting Zone:** This is formed by mapping the ZIP code from the business filing address of the EIN on Form 1120, 1120S, or 1065 to its commuting zone.

- **Value Added:** Line 3 of Form 1120 for C-Corporations, Form 1120S for S-Corporations, and Form 1065 for partnerships. Line 3 is the difference between Revenues, reported on Line 1c, and the Cost of Goods Sold, reported on Line 2. We replace non-positive value added with missing values.

  – For manufacturers (NAICS Codes beginning 31, 32, or 33) and miners (NAICS Codes beginning 212), Line 3 is equal to Value Added minus Production Wages, defined as wage compensation for workers directly involved in the production process, per Schedule A, Line 3 instructions. If we had access to data from Form 1125-A, Line 3, we could directly add back in these production wages to recover value added. Without 1125-A, Line 3, we construct a measure of Production Wages as the difference between the Wage Bill and the Firm-reported Taxable Labor Compensation, defined below, as these differ conceptually only due to the inclusion of production wages in the Wage Bill.

- **Value Added Net of Depreciation:** Value Added minus Depreciation, where Depreciation is reported on Line 20 on Form 1120 for C-corporations, Line 14 on Form 1120S for S-corporations, and Line 16c on Form 1065 for partnerships.

- **EBITD:** We follow Kline et al. (2019) in defining Earnings Before Interest, Taxes, and Depreciation (EBITD) as the difference between total income and total deductions other than interest and depreciation. Total income is reported on Line 11 on Form 1120 for C-corporations, Line 1c on Form 1120S for S-corporations, and Line 1c on Form 1065 for Partnerships. Total deductions other than interest and depreciation are computed as Line 27 minus Lines 18 and 20 on Form 1120 for C-corporations, Line 20 minus Lines 13 and 14 on Firm 1120S for S-corporations, and Line 21 minus Lines 15 and 16c on Form 1065 for partnerships.

- **Operating Profits:** We follow Kline et al. (2019), who use a similar approach to Yagan (2015), in defining Operating Profits as the sum of Lines 1c, 18, and 20, minus the sum of Lines 2 and 27 on Form 1120 for C-corporations,, the sum of Lines 1c, 13, and 15, minus the sum of Lines 2 and 20 on Form 1120S for S-corporations, and the sum of Lines 1c, 16, and 16c, minus the sum of Lines 2 and 21 on Form 1065 for partnerships.

- **Firm-reported Taxable Labor Compensation:** This is the sum of compensation of officers and salaries and wages, reported on Lines 12 and 13 on Form 1120 for C-corporations, Lines 7 and 8 on Form 1120S for S-corporations, and Lines 9 and 10 on Form 1065 for Partnerships.

- **Firm-reported Non-taxable Labor Compensation:** This is the sum of employer pension and employee benefit program contributions, reported on Lines 17 and 18 on Form 1120 for C-corporations, Lines 17 and 18 on form 1120S for S-corporations, and Lines 18 and 19 on Form 1065 for Partnerships.

- **Multinational Firm:** We define an EIN as a multinational in year t if it reports a non-zero foreign tax credit on Schedule J, Part I, Line 5a of Form 1120 or Form 1118, Schedule B, Part III, Line 6 of Form 1118 for a C-corporation in year t, or if it reports a positive Total Foreign Taxes Amount on Schedule K, Line 16l of of Form 1065 for a partnership in year t.

- **Foreign Ownership:** We define an EIN as foreign-owned in year t if it files Form 5472 in year t. The country of foreign ownership is also reported on Form 5472.

- **Tenure:** For a given TIN, we define tenure at the EIN as the number of prior years in which the EIN was the highest-paying. A TIN has No Tenure if Tenure is zero years and has High Tenure if tenure is at least five years.

- **Age and Sex:** Age at t is the difference between t and birth year reported on Data Master-1 (DM-1) from the Social Security Administration, and sex is the gender reported on DM-1 (see Chetty et al. (2011) for further details on the DM-1 link). We define Young Age as Age less than or equal to 45 years, and Not Young Age as Age greater than 45 years.

- **Gross State Lottery Winnings**: This is the total reported in Box 1 of Form W-2G when the form is identified as a state lottery payment in Box 3 of Form W-2G.

- **Adjusted Gross Income**: This is the tax-payer unit (TPU) adjusted gross income reported on Form 1040, divided by 2 in households of married filers.

- **Observed Capital Income:** This is the tax-payer unit (TPU) total observed capital income (excluding capital gains), defined as the sum of dividends, interest income, pension income, rent and royalty income, and miscellaneous Schedule E rental income. All components are reported on Form 1040. We divide by 2 in households of married filers.

- **Marginal Tax Rate:** This is the tax-payer unit (TPU) change in total taxes owed (state + federal) for a $1 change in TPU wage earnings. We calculate this using a tax calculator written by Jon Bakija.

- **Total Taxes Owed:** This is the tax-payer unit (TPU) total taxes owed (state + federal). We calculate this using a tax calculator written by Jon Bakija.

Sample Definitions:

- **Analysis Sample:** A TIN belongs to the Analysis Sample in year t if (a) her highest-paying EIN on form W-2 has positive Value Added in year t, (b) associated Earnings are at least \$15,000 in year t, (c) the commuting zone and 3-digit NAICS code of the EIN are valid in year t, and (d) the TIN is matched to SSA records and the age associated with the TIN at t is between 25 and 60.

- **Stayers Sample:** A worker belongs to the Stayers Sample in year t if (a) the worker belongs to the Analysis Sample in years t, t-1, . . . , t-7, (b) her associated highest-paying EIN is the same in years t, t-1, . . . , t-7, (c) the commuting zone and industry associated with her highest-paying EIN are the same in years t, t-1, . . . , t-7, (d) there are at least 10 stayers per firm, and (e) there are at least 10 firms per commuting zone and industry.

- **Movers Sample:** A worker belongs to the Movers Sample if (a) the worker belongs to the Analysis sample in years t and s, (b) her associated highest-paying EIN is different in years t and s, and (c) there are at least two movers associated with the EIN.

- **State Lottery Sample:** A worker belongs to the State Lottery Sample in year t if (a) she received a state lottery payment on Form W-2G between 2001 and 2016, (b) the worker is not missing age or sex data from SSA records, (c) she was 21 to 64 years old at the time of receiving the Form W-2G, and (d) her first recorded W-2G state lottery payment between 2001 and 2016 was for \$30,000 or more.

- **Procurement Auctions Sample:** A firm belongs to the Procurement Auctions Sample if the matching algorithm detects a string match on name and address. The algorithm, which uses a "fuzzy matching" approach with match quality measured using a sieve, is validated using a subsample of 5 states which provided the firm's EIN in the procurement auction records and thus permit exact matching. Table A.4(a) shows that the algorithm outperforms a simple text search (85% versus 79%).

	Goods				Services				All
	Midwest	Northeast	South	West	Midwest	Northeast	South	West	All
Panel A.	**Full Sample**								
Observation Counts:									
Number of FTE Worker-Years	42,910,324	26,701,886	40,332,913	31,598,149	69,049,669	62,399,969	103,263,800	71,385,819	447,642,529
Number of Unique FTE Workers	9,319,084	6,088,816	10,218,947	7,714,829	17,315,144	15,168,284	26,530,182	17,953,911	89,579,704
Number of Unique Firms with FTE Workers	294,907	232,740	439,823	329,721	1,051,608	1,055,084	1,908,800	1,314,677	6,479,326
Number of Unique Markets with FTE Workers	1,514	270	1,780	916	4,108	761	4,926	2,509	16,164
Group Counts:									
Mean Number of FTE Workers per Firm	22.1	17.8	16.1	16.3	10.4	9.7	9.5	9.6	11.4
Mean Number of FTE Workers per Market	2,007.0	6,778.8	1,581.7	2,524.2	1,217.4	5,623.1	1,488.4	2,084.0	1,906.6
Mean Number of Firms per Market with FTE Workers	91.0	380.6	98.0	155.2	117.0	577.9	156.2	216.3	166.9
Outcome Variables in Log $:									
Mean Log Wage for FTE Workers	10.76	10.81	10.70	10.81	10.61	10.74	10.62	10.70	10.69
Mean Value Added for FTE Workers	17.36	16.80	16.67	16.64	16.18	16.04	15.94	16.07	16.31
Firm Aggregates in $1,000:									
Wage Bill per Worker	43.6	50.7	42.2	52.9	34.3	44.2	35.8	40.3	40.9
Value Added per Worker	91.2	107.5	85.1	91.6	90.5	111.1	94.2	92.3	95.2
Panel B.	**Movers Sample**								
Observation Counts:									
Number of FTE Mover-Years	17,458,234	11,545,098	18,078,675	15,521,491	31,647,628	28,398,961	50,074,776	35,344,937	208,069,800
Number of Unique FTE Movers	4,125,425	2,830,268	4,822,238	3,877,827	7,724,643	6,663,264	11,909,494	8,324,587	32,077,850
Number of Unique Firms with FTE Movers	188,405	144,294	265,504	215,212	571,413	549,162	1,019,393	700,921	3,560,534
Number of Unique Markets with FTE Movers	1,463	266	1,753	878	3,915	755	4,783	2,359	15,609
Group Counts:									
Mean Number of FTE Movers per Firm with FTE Movers	13.5	11.9	11.2	11.6	8.2	7.9	7.9	8.2	8.9
Mean Number of Movers per Market with FTE Movers	862.4	2,964.1	730.3	1,310.7	597.7	2,617.4	759.3	1,116.4	936.7
Mean Number of Firms per Market with FTE Movers	64.0	248.9	65.3	112.8	72.6	332.3	96.1	136.8	105.0
Outcome Variables in Log $:									
Mean Log Wage for FTE Movers	10.76	10.81	10.70	10.81	10.61	10.74	10.62	10.70	10.69
Mean Value Added for FTE Movers	17.36	16.80	16.67	16.64	16.18	16.04	15.94	16.07	16.31
Panel C.	**Stayers Sample**								
Sample Counts:									
Number of 8-year Worker-Firm Stayer Spells	2,588,628	1,777,928	1,237,821	1,150,115	2,315,238	2,527,212	2,609,997	2,207,552	16,506,865
Number of Unique FTE Stayers in Firms with 10 FTE Stayers	798,575	532,507	416,549	354,518	740,091	764,699	865,629	724,155	5,217,960
Number of Unique Firms with 10 FTE Stayers	13,884	10,896	9,409	9,767	18,083	19,475	19,626	16,185	117,698
Number of Unique Markets with 10 Firms with 10 FTE Stayers	197	111	216	104	335	213	438	219	1,826
Outcome Variables in Log $:									
Mean Log Wage for FTE Stayers	10.95	10.99	10.97	10.99	10.90	11.01	10.96	11.05	10.97
Mean Log Value Added for FTE Stayers	18.04	17.56	17.46	16.56	17.45	17.23	17.89	17.93	17.61

Table A.1: Detailed sample characteristics

Notes: This table provides a detailed examination of the full sample, movers sample, and stayers sample.

B Appendix: Earnings process, value added process and pass-through

B.1 Explanation of the Empirical Approach

In this section, we use the panel data on workers and firms to describe key features of the U.S. labor market. We begin by describing the statistical model of earnings that we will apply to this data. Next, we present the empirical findings, and then discuss how they motivate and guide our choices of how to model the labor market.

B.1.1 Statistical model of earnings

We assume that workers' earnings can be described by the following equation:

$$\log W_{it} = \mathcal{X}_{it}'\vartheta + w_{it}, \tag{3}$$

where W_{it} denotes the earnings for individual i in year t, $\mathcal{X}_{it}$ is a vector of covariates which includes a full set of indicators for calendar years and a cubic polynomial in age, and w_{it} denotes log earnings net of age effects and common aggregate time trends. As described below, we allow w_{it} to depend on both the workers' own productivity and the firm in which she works. Our measure of firm performance is value added, which is determined by the equation:

$$\log Y_{jt} = \mathcal{Z}_t'\varphi + y_{jt}, \tag{4}$$

where Y_{jt} denotes the value added for firm j in year t, $\mathcal{Z}_t$ includes a full set of indicators for calendar years, and y_{jt} is log value added net of common aggregate time trends. The key elements of equations (3) and (4) are the time series properties of w_{it} and y_{jt}, which we now specify.

Specification of processes We assume that y_{jt} evolve according to the following process:

$$\begin{aligned} y_{jt} &= \zeta_j + y^p_{jt} + \xi_{jt} + \delta^y \xi_{jt-1} \\ y^p_{jt} &= y^p_{jt-1} + u_{jt}, \\ u_{jt} &= \tilde{u}_{jt} + \bar{u}_{r(j),t} \end{aligned} \tag{5}$$

where $r(j)$ denotes the market of firm j, ζ_j is a fixed effect for the firm, and the time-varying part of y_{jt} is decomposed into a permanent component, assumed to follow a unit root process with innovation shock u_{jt}, and a transitory component, which is assumed to follow a MA(1) process with coefficients δ^y and innovation variance σ^2_ξ. The permanent innovation u_{jt} consists of a common innovation to all firms in a given market r, $\bar{u}_{r(j),t} \equiv \mathbb{E}\left[u_{jt}|r(j){=}r\right]$, and an idiosyncratic innovation specific to the firm, $\tilde{u}_{jt} \equiv u_{jt}{-}\bar{u}_{r(j),t}$.

We assume that w_{it} evolve according to the following process:

$$\begin{aligned} w_{it} &= \phi_{ij(i,t)} + w^p_{it} + \nu_{it} + \delta^w \nu_{it-1} \\ w^p_{it} &= w^p_{it-1} + \gamma \tilde{u}_{j(i,t),t} + \Upsilon \bar{u}_{r(i,t),t} + \mu_{it}, \end{aligned} \tag{6}$$

where $j(i,t)$ and $r(i,t)$ denote the firm and market of worker i in year t, and ϕ_{ij} is a fixed effect for worker i if she works in firm j. The time-varying part of w_{it} is decomposed into a permanent component w^p_{it} and a transitory component, assumed to follow a MA(1) process with coefficients δ^w and innovation variance σ^2_v. The permanent earnings component evolves for three reasons: worker-specific innovations μ_{it}, pass through of firm-specific value added shocks $\gamma\tilde{u}_{j(i),t,t}$, and pass through of market level value added shocks $\Upsilon\bar{u}_{r(i,t),t}$.

Parameters of interest and assumptions Our interest is centered on two aspects of this statistical model of earnings. The first is how changes in firm performance affect the earnings of incumbent workers, as measured by the pass-through rates γ and Υ. The second is the determinants of the cross-sectional distribution of earnings, which we measure by decomposing ϕ_{ij} into components that capture worker heterogeneity, firm-specific wage premiums, worker sorting, and interactions between worker and firm effects.

For these purposes, it is necessary to invoke some restrictions on the statistical model of earnings. Let $J = \{j(i,t)\}_{i,t}$ and $U = \{\tilde{u}_{jt}, \bar{u}_{r(j),t}\}_{j,t}$ and $Q = \{\xi_{jt}\}_{j,t}$. We make the following assumptions:

Assumption 1. $\mathbb{E}[\xi_{jt}|r(j)=r, J, U] = \mathbb{E}[\xi_{jt'}\xi_{jt}|r(j)=r, J, U] = 0$ *for all* j, r, t, t'.

Assumption 2. $\mathbb{E}[\mu_{it}, \nu_{it}|J, U, Q] = 0$ *for all* i, t.

Assumption 1 is the same restriction on the error structure of the value added process as in Guiso et al. (2005). It implies that transitory shocks to value added are mean zero and uncorrelated with past transitory shocks to value added. Assumption 2 is a condition on the relationship between the worker-specific innovations to earnings, worker mobility, and innovations to firm value added. The assumption embodies two types of economic restrictions. The first restriction, from conditioning on $j(i,t)$, implies that mobility is exogenous to the worker-specific innovations to earnings (which are paid to the worker independent of the choice of firm). This is the same restriction on worker mobility as invoked in the Abowd et al. (1999) model. The second restriction, from the conditioning on the innovations to firm value added, implies that the worker-specific innovations to earnings neither co-vary across coworkers nor with shocks to firm value added. This is the same restriction as in Guiso et al. (2005).

It is important to observe what is *not* being restricted under Assumptions 1 and 2. First, we do not restrict whether or how workers sort into firms according to the worker effects, the firm effects, or the interactions between the worker and firm effects. Second, we do not restrict whether or what type of workers move across firms in response to innovations to firm value added. In fact, workers with different values of ϕ_{ij} may have arbitrarily different mobility patterns. Third, the statistical model of earnings does not specify why individuals choose the firm that they do. However, it also does not preclude the possibility that individuals choose firms to maximize earnings or utilities. For instance, Assumptions 1 and 2 are consistent with each worker choosing the firm that offers his preferred combination of wages and non-wage attributes.

B.1.2 Pass through of firm shocks

In this section, we are interested in estimating the parameters γ and Υ, which we refer to as the *pass-through rates* of firm-specific and market level value added shocks. Before presenting estimates of the pass-through rates, we show how these parameters can be identified through a difference-in-differences (DiD) strategy.

Identification, moment conditions and DiD representation To compare with existing work, we first consider a special case of the statistical model of earnings where $\gamma = \Upsilon$. That is, we assume the pass-through rate of an idiosyncratic value added shock to the current firm is of the same size as the pass-through rate of a value added shock to all firms in the current market. We focus on the sample of stayers as captured by the indicator variable $S_i = 1[j(i,1)=...=j(i,T)]$.

Assumptions 1 and 2 give the following moment conditions:

$$\mathbb{E}\left[\Delta y_{j(i)t}\left(w_{it+\tau} - w_{it-\tau'} - \gamma\left(y_{j(i),t+\tau} - y_{j(i),t-\tau'}\right)\right)|S_i{=}1\right] = 0 \quad (7)$$
$$\text{for } \tau \geq 2, \tau' \geq 3$$

Solving for γ we identify the pass through of a firm-specific shock to the earnings of incumbent workers:

$$\gamma = \frac{\mathbb{E}\left[\Delta y_{j(i)t}\left(w_{it+\tau} - w_{it-\tau'}\right)|S_i{=}1\right]}{\mathbb{E}\left[\Delta y_{j(i)t}\left(y_{j(i),t+\tau} - y_{j(i),t-\tau'}\right)|S_i{=}1\right]}$$

Thus, we can identify the pass through of a firm-specific shock from our panel data on firms and workers.

DiD interpretation

To interpret this identification result and assess the underlying assumptions, note that the statistical model of earnings includes fixed effects for time and agents. By controlling for these fixed effects we obtain a DiD strategy, looking within workers and firms while eliminating common changes over time in the labor market or the economy more generally. To see the DiD representation, suppose for simplicity the workers can be assigned to two groups of firms: one half has $\Delta y_{j(i)t} = +\delta$ and the other half has $\Delta y_{j(i)t} = -\delta$. We then get the following interpretation of γ as the ratio of two DiDs.

$$\gamma = \frac{\mathbb{E}\left[w_{it+\tau} - w_{it-\tau'}|+\delta, S_i{=}1\right] - \mathbb{E}\left[w_{it+\tau} - w_{it-\tau'}|-\delta, S_i{=}1\right]}{\mathbb{E}\left[y_{j(i),t+\tau} - y_{j(i),t-\tau'}|+\delta, S_i{=}1\right] - \mathbb{E}\left[y_{j(i),t+\tau} - y_{j(i),t-\tau'}|-\delta, S_i{=}1\right]}$$

Under an assumption of common underlying trends between the two groups, the numerator gives the treatment effect on log earnings; the denominator gives the treatment effect on log value added; and the ratio gives the elasticity of earnings with respect to value added.

Graphical evidence

In Figure A.1, we empirically assess the DiD strategy. The figure is constructed in the following way: In any given calendar year (denoted period $t = 0$), we i) order firms according to the increase $\Delta y_{j(i)t}$; ii) separate the firms at the median in the distribution of $\Delta y_{j(i)t}$, letting the upper half constitute the treatment firms and the lower half the control firms; and iii) plot the differences in y_{jt} between these two groups in period $t = 0$ as well as in the years before (periods $t < 0$) and after (periods $t > 0$). We perform these three steps separately for various calendar years, always weighting each firm by the number of workers. The solid (dashed) black line represents the difference in log value added (wages) for the treatment and control firms where each firm is weighted by the number of workers.

By construction, the treatment and control groups differ in the value added growth from period $t-1$ to period t. On average, firms in the treatment group experience about 30 percentage points larger growth in value added as compared to firms in the control group. According to the value added process (5), the growth in value added should be the sum of a permanent component and a transitory, mean-reverting component. Due to the transitory component, $\Delta y_{j(i)t}$ could be correlated with $\Delta y_{j(i)\tau}$ at $\tau = t-2, ..., t+2$. However, $\Delta y_{j(i)t}$ should be orthogonal to $\Delta y_{j(i)\tau}$ in the periods before $\tau = t-2$ and after $\tau = t+2$. Consistent with this orthogonality condition, the figure shows a very similar trend in log value added between the treatment and control group at these periods. Reassuringly, firms that experienced large growth in value added in period 0 are no more or less likely to experience large growth in value added in periods -6 to -3 or in periods 3 to 6.

The dashed black line performs the same exercise, but this time for log wages of incumbent workers who stay in the firm in all six years. On average, workers in treatment firms experience an additional 5 percentage points increase in earnings in period 0 as compared to workers in the control firms. Interpreted through the lens of the DiD design, this finding suggests a pass-through rate of firm shocks γ above .15. The growth in earnings is also the sum of a permanent component and a transitory, mean-reverting component. Therefore, Δw_{it} could be correlated with $\Delta w_{i\tau}$ at $\tau = t-2, ..., t+2$, but it should be orthogonal to $\Delta w_{i\tau}$ in the periods before $\tau = t-2$ and after $\tau = t+2$. Reassuringly, the dashed line shows a very similar trend in log earnings between workers in the treatment and control group during these periods.

Firm versus market level shocks

We now shift attention to the general case where γ may differ from Υ, thereby allowing the earnings of an incumbent worker to respond differently to an idiosyncratic value added shock to the current firm than to a (same size) shock to all firms in a given market. To identify the firm-level pass-through rate γ, we then need to demean the variables of interest using a within market times year transformation, $\tilde{w}_{it} = w_{it} - \mathbb{E}\left[w_{it}|r(i,t){=}r\right]$ and $\tilde{y}_{jt} = y_{jt} - \mathbb{E}\left[y_{jt}|r(j){=}r\right]$. Assumptions 1 and 2 then give the following moment conditions that we can use to identify the pass-through rates of firm-specific value and market level added shocks by solving for γ and Υ :

$$\mathbb{E}\left[\Delta\tilde{y}_{j(i),t}\left(\tilde{w}_{it+\tau} - \tilde{w}_{it-\tau'} - \gamma\left(\tilde{y}_{j(i),t+\tau} - \tilde{y}_{j(i),t-\tau'}\right)\right)|S_i{=}1\right] = 0 \tag{8}$$

$$\mathbb{E}\left[\Delta\bar{y}_{j(i),t}\left(\bar{w}_{it+\tau} - \bar{w}_{it-\tau'} - \Upsilon\left(\bar{y}_{j(i),t+\tau} - \bar{y}_{j(i),t-\tau'}\right)\right)|S_i{=}1\right] = 0 \tag{9}$$

$$\text{for } \tau \geq 2, \tau' \geq 3$$

where $\bar{y}_{r(j),t} \equiv \mathbb{E}\left[y_{jt}|r(j){=}r\right]$ and $\bar{w}_{r(j),t} \equiv \mathbb{E}\left[w_{jt}|r(j){=}r\right]$.

The red and blue lines in Figure A.1 represent the differences between the treatment and control group in $(\tilde{w}_{it}, \tilde{y}_{jt})$ and $(\bar{w}_{r(i,t),t}, \bar{y}_{r(j),t})$ over time. These lines are constructed in the same way as the black lines, except the red and blue lines use the demeaned variables $\tilde{w}$ and $\tilde{y}$ and the market averages $\bar{w}$ and $\bar{y}$, respectively. Comparing the red solid line to the red dashed line reveals that conditioning on the full set of year times market fixed effects attenuates slightly the treatment effect on log earnings relative to the treatment effect on log value added. Interpreted through the lens of the DiD design, this finding suggests the estimated pass-through rate of a firm-specific shock will be slightly lower once we allow for Υ to differ from γ. By way of comparison, the DiD applied to the market averages of wages and value added suggests a relatively large estimate of Υ. Thus, we expect the estimated pass-through rate of an idiosyncratic value added shock to the current firm γ to be smaller than the pass-through rate of a same size shock to all firms in the market Υ.

B.2 Tables and Figures

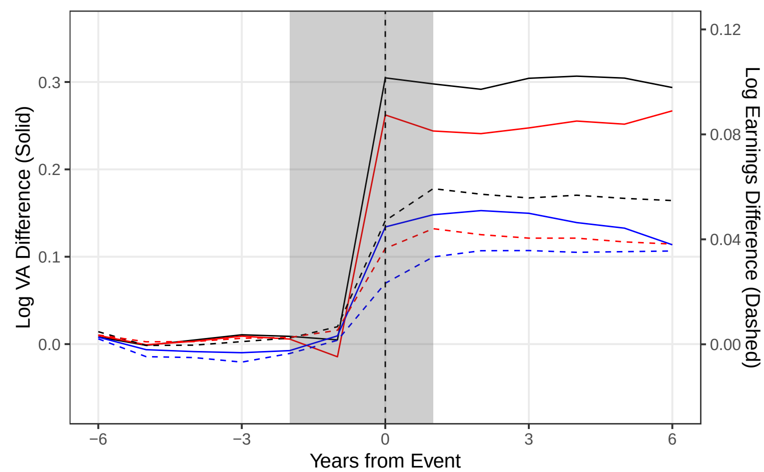
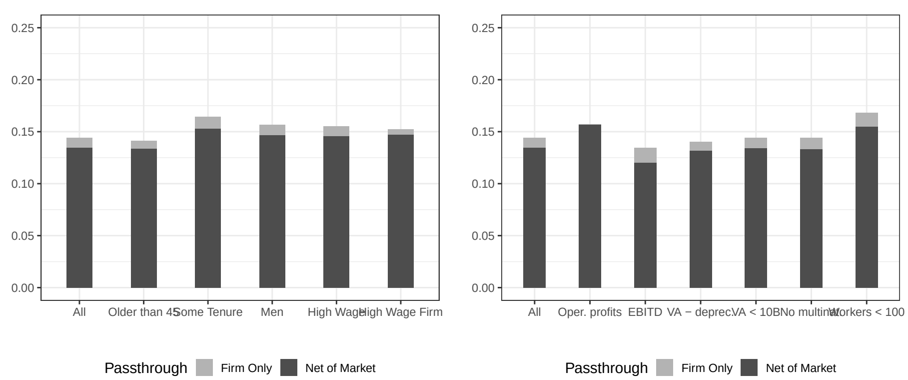


Figure A.1: Differences-in-differences representation of the estimation strategy

Notes: This figure displays the mean differences in log value added (solid lines) and log earnings (dotted lines) between firms that receive an above-median versus below-median log value added change at event time zero. Results are presented for the unconditional measures of log value added and log earnings (black lines), for the measures of log value added and log earnings net of market interacted with year effects (red lines), and for the averages of log value added and log earnings by market and year (blue lines) The shaded area denotes the time periods during which the orthogonality condition need not hold in the identification of the permanent pass-through rate.



(a) Heterogeneity across Workers

(b) Heterogeneity across Firms and V.A. Measures

Figure A.2: Pass-through Heterogeneity

Notes: This figure displays heterogeneity in the GMM estimates of the pass-through, both firm only (imposing $\Upsilon = \gamma$) and removing market by year means (permitting $\Upsilon \neq \gamma$).

	GMM Estimates of Joint Process			
	Firm Only		Accounting for Markets	
	Log Value Added	Log Earnings	Log Value Added	Log Earnings
Panel A.	**Process: MA(1)**			
Total Growth (Std. Dev.)	0.31	0.17	0.29	0.16
	(0.01)	(0.00)	(0.01)	(0.00)
Permanent Shock (Std. Dev.)	0.20	0.10	0.17	0.10
	(0.01)	(0.00)	(0.01)	(0.00)
Transitory Shock (Std. Dev.)	0.18	0.10	0.17	0.10
	(0.01)	(0.00)	(0.01)	(0.00)
MA Coefficient, Lag 1	0.09	0.15	0.09	0.15
	(0.01)	(0.00)	(0.01)	(0.00)
MA Coefficient, Lag 2	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Permanent Passthrough Coefficient		0.14		0.13
		(0.01)		(0.01)
Transitory Passthrough Coefficient		-0.01		0.00
		(0.01)		(0.00)
Market Passthrough Coefficient				0.18
				(0.02)
Panel B.	**Process: MA(2)**			
Total Growth (Std. Dev.)	0.31	0.17	0.29	0.16
	(0.01)	(0.00)	(0.01)	(0.00)
Permanent Shock (Std. Dev.)	0.20	0.10	0.17	0.10
	(0.01)	(0.00)	(0.00)	(0.00)
Transitory Shock (Std. Dev.)	0.17	0.10	0.17	0.10
	(0.01)	(0.00)	(0.01)	(0.00)
MA Coefficient, Lag 1	0.05	0.21	0.07	0.21
	(0.05)	(0.01)	(0.04)	(0.01)
MA Coefficient, Lag 2	-0.03	0.04	-0.01	0.04
	(0.03)	(0.00)	(0.02)	(0.00)
Permanent Passthrough Coefficient		0.15		0.13
		(0.01)		(0.01)
Transitory Passthrough Coefficient		-0.02		0.00
		(0.01)		(0.00)
Market Passthrough Coefficient				0.18
				(0.03)

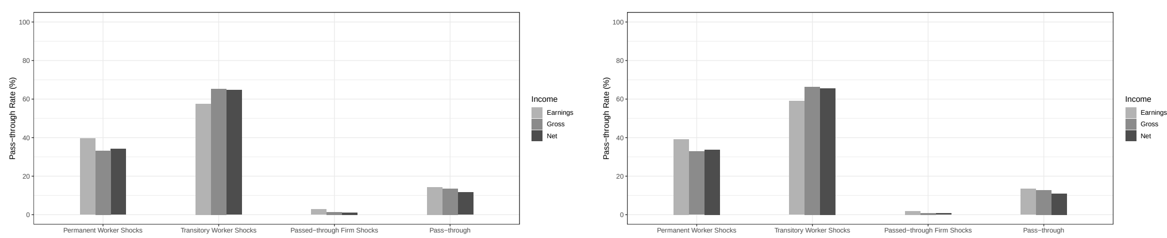
Table A.2: Estimated Process for Log Earnings and Pass-through

Notes: This table displays the parameter estimates of the log value added and log earnings growth processes as well as the passthrough coefficients when using the GMM estimator. It presents these estimates for the firm only model (which imposes $\Upsilon = \gamma$) as well as the model in which firm and market pass-through coefficients are allowed to differ (which permits $\Upsilon \neq \gamma$).

	Goods				Services			
	Midwest	Northeast	South	West	Midwest	Northeast	South	West
Log Earnings:								
Unconditional	0.156	0.126	0.143	0.176	0.136	0.113	0.143	0.146
Net of Market	0.156	0.132	0.136	0.168	0.124	0.110	0.140	0.119
Market	0.157	0.101	0.163	0.211	0.205	0.139	0.154	0.264
Selection Coefficient	0.250	0.190	0.243	0.184	0.143	0.115	0.238	0.182
Log Size:								
Unconditional	0.454	0.480	0.348	0.441	0.501	0.354	0.479	0.464
Net of Market	0.492	0.517	0.352	0.455	0.480	0.360	0.585	0.444
Other Moments:								
Market Size (millions)	42.9	26.7	40.3	31.6	69.0	62.4	103.2	71.4
Labor Share	0.630	0.663	0.746	0.790	0.660	0.659	0.700	0.662
Profits per FTE Worker ($1,000)	47.6	56.9	43.0	38.7	56.5	66.9	58.4	52.0
Wagebill per FTE Worker ($1,000)	43.6	50.7	42.2	52.9	34.1	44.2	35.8	40.3

Table A.3: Detailed Passthrough Estimates and Aggregate Statistics across Regions and Sectors

Notes: In this table, we present passthrough estimates for various outcomes as well as other empirical moments across broad regions and sectors.

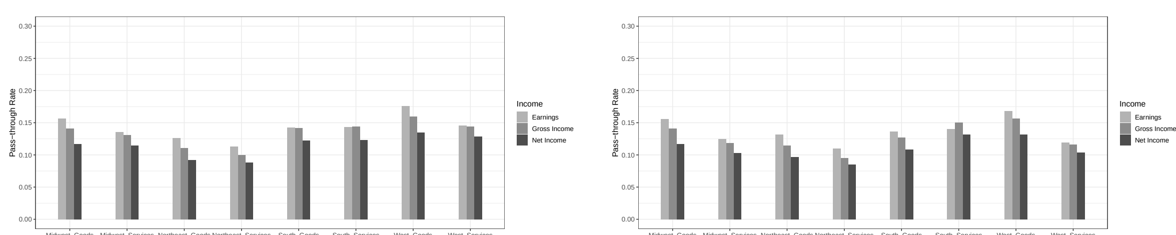


(a) Firm Only

(b) Net of Market

Figure A.3: Growth and Pass-through Estimation for Alternate Income Concepts

Notes: In subfigure (a), we present the shares of (i) earnings, (ii) gross income, and (iii) net income growth attributable to permanent and transitory shocks to workers and permanent shocks to firms, as well as the passthrough rate from firms to workers, in the baseline specification ("Firm Only"). In subfigure (b), we repeat this exercise when conditioning on a full set of market-year indicators ("Net of Market").



(a) Firm Only

(b) Net of Market

Figure A.4: Broad Market Heterogeneity in Passthrough Estimation by Income Measure

Notes: In this figure, we present broad market heterogeneity in the firm only and net passthrough rate of permanent shocks from firms to workers.

	Simple Search	Fuzzy Match			
	(1)	(2)	(3)	(4)	(5)
% Bidders Matched to Any Tax Record	80.2	99.9	97.6	99.9	95.8
% Bidders Matched to the True Tax Record	65.3	63.0	62.5	71.0	70.3
% Potential Matches Correctly Matched to Tax Records	78.6	75.8	75.1	85.4	84.5
Algorithm Parameters:					
Match must be perfect (string score = 1.0)	✓	✗	✗	✗	✗
Match must be high-quality (string score ≥ 0.6)	✗	✗	✓	✗	✓
Prefer matches in same state as auction	✓	✗	✗	✓	✓

(a) Algorithm Match Performance

	DOT Auction Records		**Final Sample: Matched Auction-Tax Data**		
State	**Data Source**	**Includes EIN**	**Bidders in 2010**	**Share of 2010 Construction Sector:**	
			(Num. Firms)	**Value Added**	**FTE Workers**
AL	State Website	✗	196	15.7%	17.4%
AR	State Website	✗	149	7.9%	12.8%
AZ	No	✗	*	*	*
CA	State Website	✗	1,041	8.3%	11.2%
CO	FOIA Request	✓	241	12.6%	14.7%
CT	FOIA Request	✗	126	9.4%	15.5%
FL	State Website	✓	344	30.7%	10.6%
GA	BidX Website	✗	137	4.3%	7.0%
IA	BidX Website	✗	256	15.4%	20.7%
ID	BidX Website	✗	112	17.2%	13.6%
IL	No	✗	*	*	*
IN	State Website	✓	213	10.6%	16.6%
KS	BidX Website	✓	130	13.7%	21.6%
KY	No	✗	*	*	*
LA	BidX Website	✗	167	11.5%	10.8%
MA	No	✗	*	*	*
MD	No	✗	*	*	*
ME	BidX Website	✗	141	13.7%	16.9%
MI	BidX Website	✗	391	9.5%	16.3%
MN	BidX Website	✗	262	13.5%	19.8%
MO	BidX Website	✗	179	14.9%	13.3%
MS	No	✗	*	*	*
MT	FOIA Request	✗	122	15.0%	23.6%
NC	BidX Website	✗	135	5.2%	9.8%
ND	FOIA Request	✗	*	*	*
NE	No	✗	*	*	*
NH	No	✗	*	*	*
NJ	No	✗	*	*	*
NM	BidX Website	✗	*	*	*
NV	No	✗	*	*	*
NY	No	✗	*	*	*
OH	BidX Website	✗	320	43.7%	17.5%
OK	No	✗	*	*	*
OR	No	✗	*	*	*
PA	No	✗	*	*	*
SC	No	✗	*	*	*
SD	No	✗	*	*	*
TN	BidX Website	✗	140	5.3%	11.5%
TX	FOIA Request	✓	551	4.9%	9.6%
UT	No	✗	*	*	*
VA	BidX Website	✗	241	14.2%	12.0%
VT	BidX Website	✗	*	*	*
WA	BidX Website	✗	200	7.5%	14.0%
WI	BidX Website	✗	194	12.1%	14.6%
WV	BidX and State Websites	✓	103	13.7%	19.0%
National			6,792	10.7%	9.9%

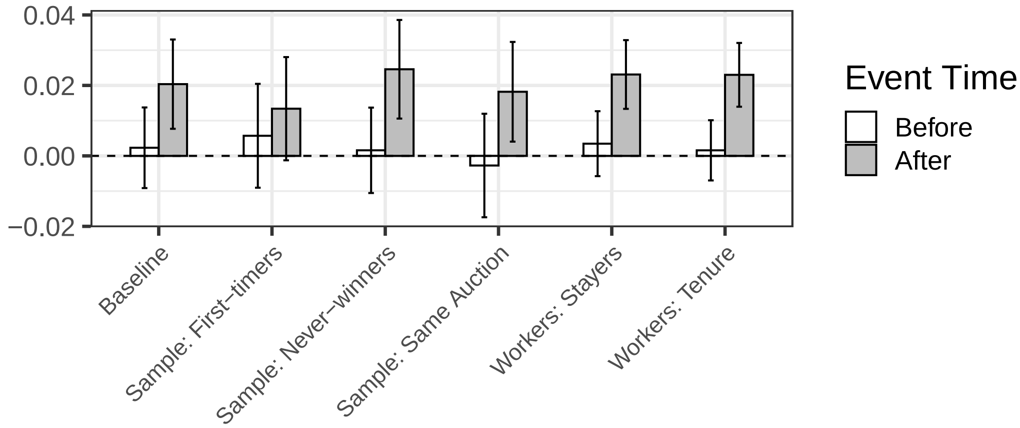
(b) Shares by State of the Matched Firms in 2010

	Sample Size		Share of the Construction Sector
Number of Firms	7,876		0.9%
Workers per Firm	46		11.7%
	Value Per Firm ($ millions)	**Mean of the Log**	**Share of the Construction Sector (%)**
Sales	19.927	15.061	12.1%
EBITD	9.159	14.075	9.6%
Intermediate Costs	14.661	14.719	12.4%
Wage bill	2.737	13.549	13.4%

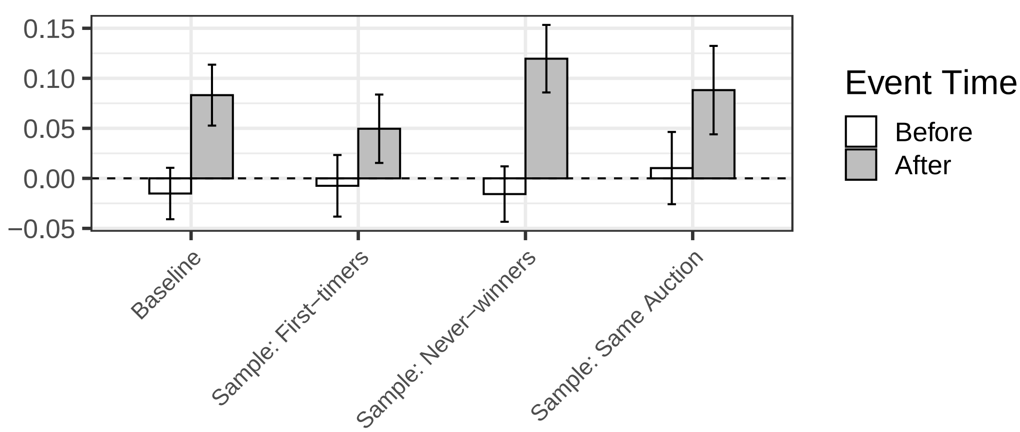
(c) Characteristics of the Matched Firms in 2010

Table A.4: Characteristics of the Procurement Auctions Sample

Notes: These table characterizes the matched sample of firms that bid in procurement auctions. Table (a) provides information on the performance of various matching algorithms. Table (b) provides information on the representativeness of the matched sample in 2010, where states with more than zero but fewer than 100 matched firms are omitted (denoted by *). Table (c) displays sample statistics (in 2010 unless stated otherwise).



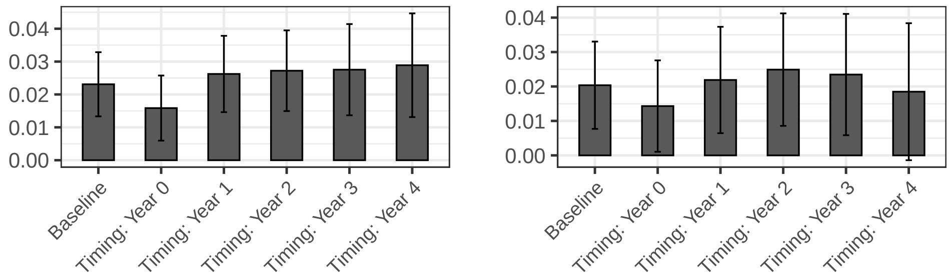
(a) Log earnings per worker (numerator of θ_{IV})



(b) Log number of employees (denominator of θ_{IV})

Figure A.5: Pass-through of Observable Demand Shocks

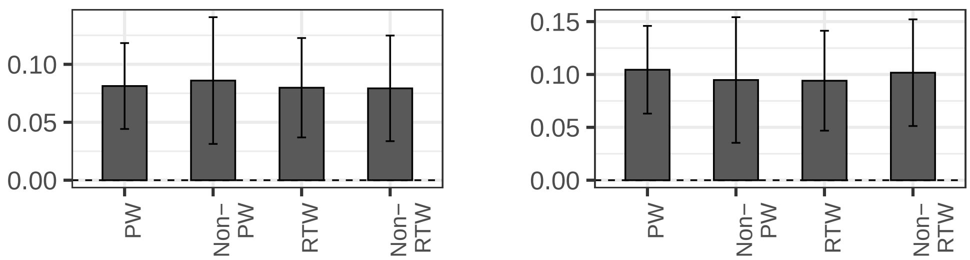
Notes: This figure presents average treatment effect on the treated estimates using the difference-in-differences specification defined in the text. All results include firm fixed effects. The control units are those firms that place a bid in a procurement auction in the same year that the reference treatment cohort wins. The omitted relative time is -2. 90% confidence intervals are displayed, clustering on firm.



(a) Log earnings per stayer

(b) Log number of employees

Figure A.6: Reduced Form Estimates at Annual Frequency



(a) Log number of employees

(b) Log wage bill

Figure A.7: Reduced Form Estimates by Right-to-Work or Prevailing Wage States

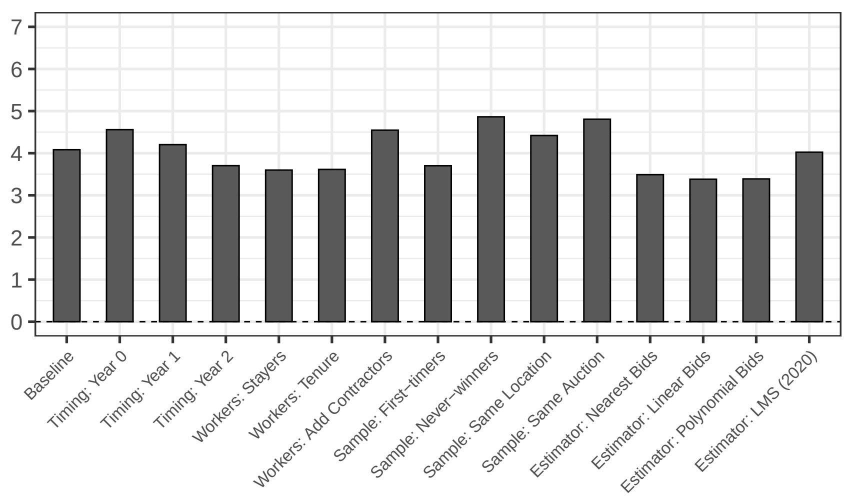
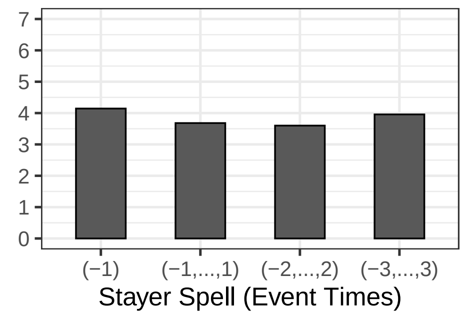
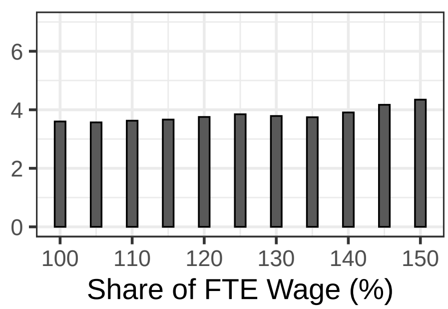


Figure A.8: Labor Supply Elasticity: Main Estimates and Robustness Checks

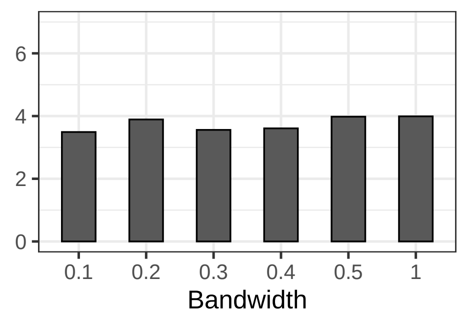


(a) Robustness by Stayer Spell

(b) Robustness by Tenure Length



(c) FTE Threshold



(d) Bandwidth Threshold

Figure A.9: Labor Supply Elasticity: Baseline Estimate and Alternative Specifications

Outcome Sample	First Stage (Std. Error)	Reduced Form (Std. Error)	Second Stage (Std. Error)
Procurement auction shock at firm-level			
8,677 unique auction bidders	0.143 (0.039)	0.020 (0.006)	0.142 (0.068)
Shift-share industry value added shock			
667 unique commuting zones	0.708 (0.216)	0.134 (0.061)	0.189 (0.041)

Table A.5: Additional details regarding passthrough estimation

Notes: This table provides additional details on the passthrough estimation using procurement auctions and shift-share instruments.

C Appendix: Movers Analyses

C.1 Limited mobility bias

Even if the restrictions discussed in the text hold, it is challenging to draw inference about the inequality contribution from firm effects and worker sorting. A key challenge is the incidental parameter bias caused by the large number of firm-specific parameters that are solely identified from workers who move across firms. The analysis of Andrews et al. (2008) suggests this limited mobility bias can be substantial. With few movers per firm, the firm component is biased upwards while the sorting component is biased downwards, with the size of the bias depending inversely on the degree of worker mobility among firms.

To get a better sense of the scope for limited mobility bias in the U.S. data, we would ideally apply the AKM estimator to alternative samples of workers and firms that are comparable except for the number of movers per firm. Figure A.10 presents the results from such an analysis, suggesting that the variance of firm effects declines monotonically as the number of movers per firm increases. To construct this figure, we consider a subsample of firms with reasonably many movers; that is, at least 15 movers per firm over the period 2001-2008. Applying AKM to this subsample gives an estimate of the variance of firm effects of 6.7 percent. Next, we remove movers randomly within firms (keeping the connected set of firms approximately the same) before re-estimating the AKM model. The solid line displays the AKM estimates of the variance of firm effects after randomly removing movers. Consistent with limited mobility bias, the fewer the number of movers per firm, the larger the variance of firm effects. For approximately the same set of firms, the estimated variance of firm effects is several times as large (23 percent) if we only keep ten percent of the movers within each firm (on average, 7 movers per firm) as compared to what we obtained if we keep all the movers per firm (at a minimum 15 and, on average, 62 movers per firm). By way of comparison, there are around 18 movers per firm in the full estimation sample (which roughly corresponds to the number of movers per firm when randomly removing 40% of movers).

Until recently, the procedures for addressing limited mobility bias required strong and questionable assumptions about the covariance structure of the time-varying errors (see e.g. the discussion in Card et al., 2018). To address this shortcoming, BLM and Kline et al. (2020) propose approaches to address limited mobility bias that rely on a different or weaker set of assumptions.[1] The first approach reduces the dimension of firm heterogeneity to a finite number of types. BLM show how this approach can be used to alleviate the biases arising from low mobility rates. The second approach uses a version of the Jackknife method. Kline et al. (2020) show how this approach allows one to relax the homoskedasticity assumption in the bias correction procedure proposed by Andrews et al. (2008). Since it is computationally infeasible to apply Andrews et al. (2008) and Kline et al. (2020) to very large data sets (as one needs to compute the trace of the inverse of the mobility matrix), our main analysis is based on the approach of BLM. As a robustness check, however, we use a subset of the U.S. states to assess

[1] Another possibility is to change the definition of a firm effect. See Borovickova and Shimer (2017) for such an approach.

the sensitivity of the results to the choice of procedure for addressing limited mobility bias.

In Figure A.10, the dotted line shows estimates of the variance of firm effects based on the procedure of BLM that addresses limited mobility bias. Firms are first classified into groups based on the empirical earnings distribution using the k-means clustering algorithm. The k-means classification groups together firms whose earnings distribution is most similar. Then, in a second step, the worker effects and firm effects are estimated. While the specification of BLM in Figure A.10 assumes there exists 10 firm types, Appendix Figure A.11 shows the BLM estimates do not materially change if we instead allow for 20, 30, 40 or 50 firm types. Consistent with limited mobility bias, the BLM estimates are noticeably smaller than the standard AKM estimates in the samples with few movers. As expected, the AKM estimates become more similar to the BLM estimates when there is a large number of movers per firm, and thus, limited mobility bias should be small.

C.2 Extensions to the AKM Model

The assumptions that $\phi_{ij} = x_i + \psi_j$ and $\gamma = \Upsilon = 0$ implies strong restrictions on the wage structure. The absence of interactions between worker and firm effects rules out strong complementaries in production, as in Shimer and Smith (2000) and Eeckhout and Kircher (2011). The assumption of no pass through of firm and market shocks is at odds with our data and a large body of evidence from many other developed countries. Thus, investigating these assumptions seems important to draw credible conclusions about the functioning of the U.S. labor market.

Non-additivity and complementarities

The assumption that $\phi_{ij} = x_i + \psi_j$ implies that all workers who move from firm j to j' will experience an earnings change of $\psi_{j'}$ - ψ_j, no matter their quality x_i. An informal way to assess this log additive structure is to perform an event study of the earnings changes experienced by workers moving between different types of firms. Card et al. (2013b) and Card et al. (2018) use matched employer-employee data from Germany and Portugal to perform such event-study analyses of the earnings changes experienced by workers moving between different types of firms. In Appendix Figure A.15, we perform the same exercise, but this time for our U.S. data. This analysis uses the movers sample. As in Card et al. (2013b) and Card et al. (2018), we define firm groups based on the average pay of coworkers.

The results from the event study mirror those reported in Card et al. (2013b) and Card et al. (2018). Workers who move to firms with more highly-paid coworkers experience earnings raises, while those who move in the opposite direction experience earnings decreases of similar magnitude. Additionally, the gains and losses for movers in opposite directions between any two groups of firms are relatively symmetric. By comparison, earnings do not change materially when workers move between firms with similarly paid coworkers. Another relevant finding from the event study is that the earnings profiles of the various groups are all relatively stable in the years before and after a job move. This lends support to Assumption 2, as it suggests that worker mobility does not seem to depend strongly on the trends in earnings beforehand or afterwards. Lastly, it is interesting to observe that the gains and losses for movers seem to be permanent. In contrast, in a large class of search models with job ladders, moves to firms that currently pay less is rationalized by arguing that these firms will pay more in the future.

Although the event study results are consistent with the log additive functional form, we cannot rule out interaction effects between worker and firm effects. Indeed, Bonhomme et al. (2019) point out that even if the functional form is non-additive, the gains and losses may look symmetric if workers making upward moves are of the same quality as those making downward moves. More generally, the degree of asymmetry one observes in the event study depends both on the magnitudes of any interaction effects and on the extent to which workers making upward moves differ in quality from those making downward moves. Thus, the event study analysis needs to be interpreted with caution.

To obtain an actual estimate of the importance of interactions between worker and firm effects, we follow BLM in using the following model of earnings:

$$w_{it} = \underbrace{\theta_{j(i,t)} \cdot x_i}_{\text{interaction}} + \psi_{j(i,t)} + \epsilon_{it} \tag{10}$$

which reduces to AKM when θ_j is the same for all firms. Under Assumptions 1 and 2 , we obtain:

$$\begin{aligned}
\mathbb{E}[w_{it+1}|j_2 \to j_1] - \mathbb{E}[w_{it}|j_1 \to j_2] &= \theta_{j_1} \left(\mathbb{E}\left[x_i|j_2 \to j_1\right] - \mathbb{E}\left[x_i|j_1 \to j_2\right]\right) \\
\mathbb{E}[w_{it+1}|j_1 \to j_2] - \mathbb{E}[w_{it}|j_2 \to j_1] &= -\theta_{j_2} \left(\mathbb{E}\left[x_i|j_2 \to j_1\right] - \mathbb{E}\left[x_i|j_1 \to j_2\right]\right)
\end{aligned}$$

where $j_1 \to j_2$ $(j_2 \to j_1)$ is an indicator for a worker moving from firm 1 to 2 (firm 2 to 1). As long as the workers moving from 1 to 2 are not exactly the same as those moving from 1 to 2, the right hand side of these equalities are non-zero and we can recover θ_{j1}/θ_{j2} from the moment condition:

$$\frac{\mathbb{E}[w_{it+1}|j_2 \to j_1] - \mathbb{E}[w_{it}|j_1 \to j_2]}{\mathbb{E}[w_{it}|j_2 \to j_1] - \mathbb{E}[w_{it+1}|j_1 \to j_2]} = \frac{\theta_{j1}}{\theta_{j2}} \tag{11}$$

Thus, provided that the composition of movers differs across firms, it is possible to identify θ_j (up to scale) for every firm. To take (10) to the data, however, it is useful to reduce the number of parameters to estimate. As above, we follow BLM in classifying firms to ten types according to the empirical earnings distribution within firms. Then we restrict θ_j to be the same for all firms of a given type.

Figure A.12 displays the estimated nonlinearities. We plot the means of log earnings for each firm type and at 10 deciles of worker heterogeneity. On the x-axis, firm types are ordered in ascending order, where "lower" and "higher" types refer to low and high mean log earnings. The results show clear evidence of worker heterogeneity: For the same type of firm, better workers earn significantly more. For a given worker, there is also some variation in log earnings between firm types, although to a lesser extent. As shown in equation (11), the parameters governing nonlinearities are identified from comparing the gains from moving from a low to a high type of firm for workers of different quality. As evident from Figure A.12, the gains from such a move are considerably larger for better workers. For example, moving from the lowest to the highest type of firm increases earnings by 22, 47 and 78 percentage points for individuals at the 20, 50 and 80 percentile in the worker quality distribution. As an alternative, we apply another estimator suggested by BLM which partitions workers into 5 discrete types, then estimates an interaction parameter for each firm-worker-type pair. The results are displayed in A.14, finding a similar pattern in which there are stronger interactions among high-type workers and high-type firms.

The evidence of nonlinearities raises several questions. To what extent do interaction effects bias the estimates from the log additive model? Are nonlinearities empirically important as a source of earnings inequality? In Table A.8, we investigate these questions by extending the AKM decomposition to incorporate the contribution from interactions between worker and firm effects. Re-arranging equation (10), we get

$$w_{it} = \underbrace{\bar{\theta}(x_i - \bar{x})}_{\tilde{x}_i} + \underbrace{\left(\psi_{j(i,t)} + \theta_{j(i,t)}\bar{x}\right)}_{\tilde{\psi}_{j(i,t)}} + \underbrace{(\theta_{j(i,t)} - \bar{\theta})(x_i - \bar{x})}_{\varrho_{ij(i,t)}} + \epsilon_{it} \tag{12}$$

where $\bar{\theta} \equiv \mathbb{E}\left[\theta_{j(i,t)}\right]$ and $\bar{x} \equiv \mathbb{E}\left[x_i\right]$. This equation decomposes the earnings of worker i in period t into three distinct components: $\tilde{x}_i$ gives the direct effect of the quality of worker i (evaluated at the average firm), $\tilde{\psi}_{j(i,t)}$ represents the direct effect of firm j (evaluated at the average worker), and $\varrho_{ij(i,t)}$ captures the interaction effect between firm j and worker i quality.

Using equation (12), we obtain a new variance decomposition of log earnings:

$$\begin{aligned} Var(w_{it}) =& Var\left[\tilde{x}_i\right] + Var\left[\tilde{\psi}_{j(i,t)}\right] + 2Cov\left[\tilde{x}_i, \tilde{\psi}_{j(i,t)}\right] \\ &+ Var\left[\varrho_{ij(i,t)}\right] + 2Cov\left[\tilde{x}_i + \tilde{\psi}_{j(i,t)}, \varrho_{ij(i,t)}\right] \end{aligned} \tag{13}$$

The first three components are informative about the inequality contribution from worker effects, firm effects and worker sorting, net of interaction effects. The last two components are informative about the inequality contribution from interaction effects, as measured by the dispersion of $\varrho_{ij(i,t)}$ across firms and the extent to $\varrho_{ij(i,t)}$ is larger in firms with high wages. If $\theta_j = \bar{\theta}$ for every firm j, then these two components would be zero, and the decomposition in (13) reduces to the standard AKM decomposition.

The results from the decomposition in (13) are presented in column (2) of Table A.8. Our estimates suggest the dispersion of interaction effects across firms explains three percent of the earnings inequality. However, the total contribution to earnings inequality from nonlinearities is muted by the interaction effects being larger in firms with higher paid workers. We also find that omitting interaction effects causes a downward bias in the firm effects and an upward bias in the worker effects.

Pass through of shocks and time-varying types

The assumption that $\gamma = \Upsilon = 0$ restricts firm effects to be constant over time. However, the significant pass-through rates imply that firm effects actually evolve over time as employers experience changes in the value added at the firm or market level. To capture this, we now let γ differ from Υ and propose an adjustment to the AKM model which allows us to isolate the time-invariant component of the firm effects.

Our approach proceeds in two steps. First, we construct an adjusted earnings measure by removing the time-varying firm and market specific component of earnings. To do so, we use the firm and market level value added multiplied by the estimated passthrough coefficients at the firm and market level (see Table 2). Second, we recover the time-invariant firm and worker effects by applying the methods of AKM or BLM to the adjusted measure of earnings. Consider the following adjusted two-way specification for earnings of workers across firms:

$$\mathbb{E}[w_{it} - \gamma(y_{j(i,t),t} - y_{j(i,t),1}) - (\Upsilon - \gamma)\left(\bar{y}_{r(i,t),t} - \bar{y}_{r(i,t),1}\right)|j(i,1),...,j(i,T)] = x_i + \psi_j.$$

The left-hand side removes the earnings dynamics due to passthrough of firm-specific shocks, $\gamma(y_{j(i,t),t} - \bar{y}_{r(i,t),t})$, and market shocks, $\Upsilon\bar{y}_{r(i,t),t}$. What remains is the worker effect x_i and the time-invariant firm effect ψ_j, which can be estimated by applying AKM or BLM to the adjusted earnings measure.

In column (3) of Table A.8, we extend the BLM decomposition of the variance of log earnings in (2) to incorporate the contribution from time-invariant and time-varying firm effects. We find that time-varying firm effects explain little if any of the variation in log earnings, and that the importance of firm effects and worker sorting do not change materially if we take the pass through of firm shocks into account. Comparing the results in column (4) to those presented in column (2) shows that time variation also has little to no explanatory power when accounting for nonlinearities.

C.3 Additional robustness checks

In our main analysis, we follow the literature in looking at individuals aged 25-60 for whom earnings exceed the full-time equivalence. This raises the question of how sensitive the results are to changing these sample selection criteria. In Appendix Figure A.16, we re-estimate the AKM model with alternative employment definitions, reporting the variance of log earnings and the firm effects. This figure also examines how the results change if we include workers aged 20-25. As expected, the variance of log earnings and the estimated firm effects increase if we include individuals earning less than the full-time equivalence. By comparison, the inclusion of younger workers do not materially change the estimates. In Appendix Figure A.18, we examine how the firm effects differ between men and women when using the AKM model. We find fairly similar patterns, with men tending to sort more strong towards firms with higher firm effects.

We repeat many of the movers analyses for the "Broad Sample". This sample is similar to the Baseline Sample considered elsewhere, but with five differences. First, firms are not required to have positive value added. Second, firms are not required to have at least two movers. Third, locations are taken from worker rather than firm forms. Fourth, we consider shorter time intervals of 2, 3, or 6 years rather than 8 years. Fifth, because the sample is smaller, we can feasibly compare many alternate estimation strategies.

Characteristics of the Broad Sample are displayed in Table A.9(a). We see that, when using a smaller numbers of years, there are fewer movers relative to the number of stayers, on average as well as across the distribution of movers. This results in a much smaller share of firms belonging to the connected set. Table A.9(a) also characterizes the leave-one-out set of Kline et al. (2020), which requires that firms are connected by at least one mover even after dropping a mover from the sample and is required to use the bias correction method of Kline et al. (2020).

The main national results for the Broad Sample are displayed in Appendix Tables A.9(b) for the connected set and A.9(c) for the leave-one-out set. In the connected set, the AKM estimator (FE) estimates a variance of firm effects of 12% to 16%, which are greater than the estimate of 9% in the Baseline Sample, and a sorting share of -12% to 1%, which are less than 5% in the Baseline Sample, which is consistent with greater bias if there are fewer movers per firm. Imposing the leave-one-out set, the FE estimates become more similar to the Baseline Sample, as the number of movers per firm rises. Applying three types of bias correction procedures by Bonhomme et al. (2019) (CRE), Andrews et al. (2008) (FE-HO), and Kline et al. (2020) (FE-HE), we find that the variance of firm effects ranges from 4% to 6%, which is similar to the 3% bias-corrected estimate in the Baseline Sample.

Using the Broad Sample, we repeat the various exercises presented above. Appendix Figure A.15 compares the event study of earnings changes for workers who move across firms in the Baseline Sample (subfigure a) and the Broad Sample (subfigure b), finding strong similarities. Appendix Figure A.11 compares BLM and CRE estimates as the number of clusters increases from 10 to 50 for the Baseline Sample (subfigure a) and the Broad Sample (subfigure b), finding that the number of clusters does not affect the estimates. Using the Broad Sample to investigate the full-time earnings threshold, we find in Appendix Figure A.19(a-b) a similar pattern as we found in Appendix Figure A.16 for the Baseline Sample. In Appendix Figure A.19(c-d), we examine the sensitivity of the estimates to the minimum number of workers per firm, finding that much of the bias in the FE estimator dissipates if only considering larger firms. Using the Broad Sample to investigate limited mobility bias by share of movers kept, we find in Appendix Figure A.19(e-h) a similar pattern as we found in Appendix Figure A.10 for the Baseline Sample.

We consider several additional robustness checks that make use of the Broad Sample. Appendix Figure A.19(i-j) shows that the qualitative results for the Baseline Sample and the Broad Sample are present even when considering very short 2-year panels. Appendix Figure A.19(k-l) demonstrates that the bias in the AKM estimator would become even stronger if we used a strict definition of movers in which we require workers to be employed for three consecutive years at each firm. Appendix Figure A.19(m-n) considers the 20 smallest states, showing that the main qualitative results hold at the state level, while Appendix Figure A.19(o-p) shows that these results hold at the state level even when using the exact solution to the estimator rather than the approximation method required for feasible estimation on large samples.

C.4 Tables and Figures

Sample:	Full Sample	≥ 2 Movers	Connected Set
Workers in 2001-2008:			
Worker-Years (Millions)	245.0	227.8	227.4
	(100.0%)	(93.0%)	(92.8%)
Unique Workers (Millions)	66.2	61.8	61.7
	(100.0%)	(93.3%)	(93.2%)
Workers in 2008-2015:			
Worker-Years (Millions)	232.9	212.4	211.9
	(100.0%)	(91.2%)	(91.0%)
Unique Workers (Millions)	64.0	58.8	58.6
	(100.0%)	(91.9%)	(91.7%)

Table A.6: Floor on Number of Movers and the Connected Set

Notes: This table demonstrates the fraction of workers lost from the sample in the AKM and BLM analysis when imposing that a firm must have at least two movers and must belong to the connected set of firms.

Years:	**2001-2008**	**2008-2015**	**Pooled**
Panel A.		**Levels**	
Total SD	0.67	0.68	0.67
Worker Effects SD	0.57	0.58	0.57
Firm Effects SD	0.20	0.20	0.20
Covariates SD	0.11	0.12	0.11
Correlation: x_i and $\psi_{j(i)}$	0.09	0.11	0.10
Correlation: x_i and $X_i'b$	0.00	0.00	0.00
Correlation: $X_i'b$ and $\psi_{j(i)}$	0.02	0.03	0.03
Panel B.		**Percentages**	
$Var(x_i + X_i'b)$	75.4%	75.5%	75.4%
$Var(x_i)$	72.9%	72.4%	72.6%
$Var(X_i'b)$	2.5%	3.1%	2.8%
$2Cov(x_i, X_i'b)$	0.0%	0.0%	0.0%
$Var(\psi_{j(i)})$	8.8%	9.0%	8.9%
$2Cov(x_i + X_i'b, \psi_{j(i)})$	4.9%	5.7%	5.3%
$2Cov(x_i, \psi_{j(i)})$	4.6%	5.4%	5.0%
$2Cov(X_i'b, \psi_{j(i)})$	0.2%	0.3%	0.3%
Residual	11.0%	9.9%	10.4%

Table A.7: Detailed AKM Decomposition

Notes: This table decomposes the variance of log earnings into components of worker effect variance, firm effect variance, the variance of time-varying observables, and the covariances among these components. Results are presented for the AKM estimator for the 2001-2008 and 2008-2015 samples, as well as their average (Pooled).

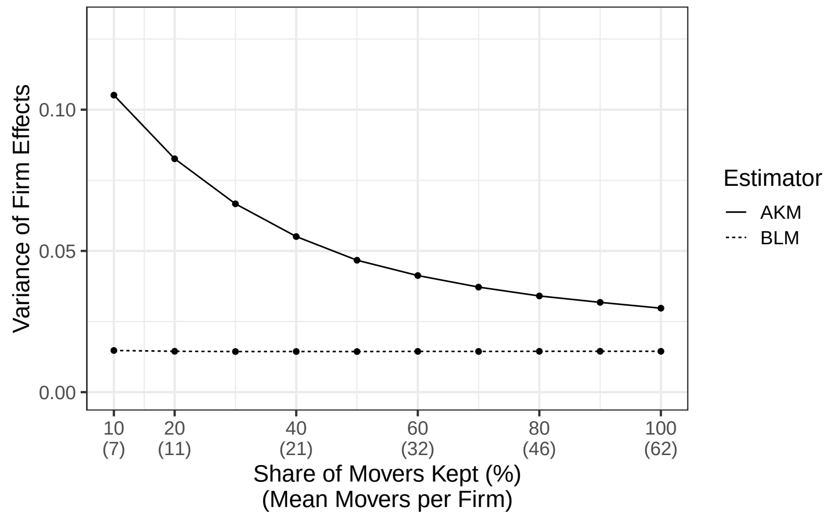
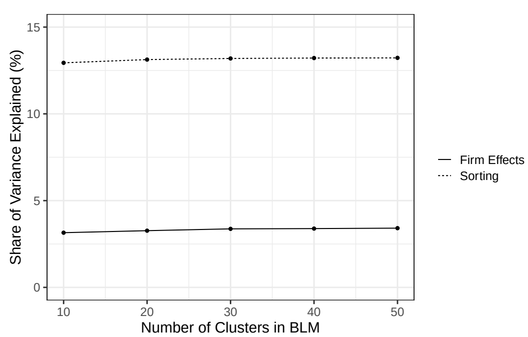
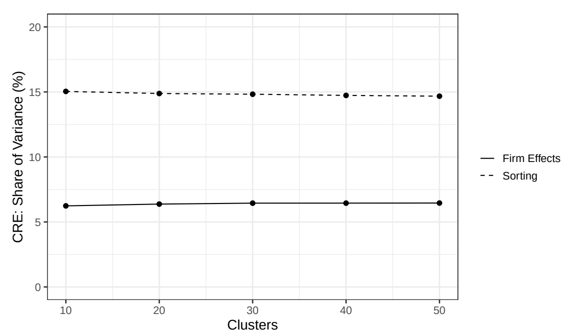


Figure A.10: Empirical Characterization of Limited Mobility Bias

Notes: In this figure, we consider the subset of firms with at least 15 movers. We randomly remove movers within each firm and re-estimate the variance of firm effects using the AKM and BLM estimators. For each estimator, we repeat this procedure several times, and then take averages of the variance estimates across these repetitions. The procedure allows us to keep the connected set of firms approximately the same and examine the bias that results from having fewer movers available in estimation.



(a) Baseline Sample



(b) Broad Sample

Figure A.11: BLM Decomposition by Number of Clusters

Notes: In this figure, we estimate the BLM decomposition for different numbers of firm clusters.

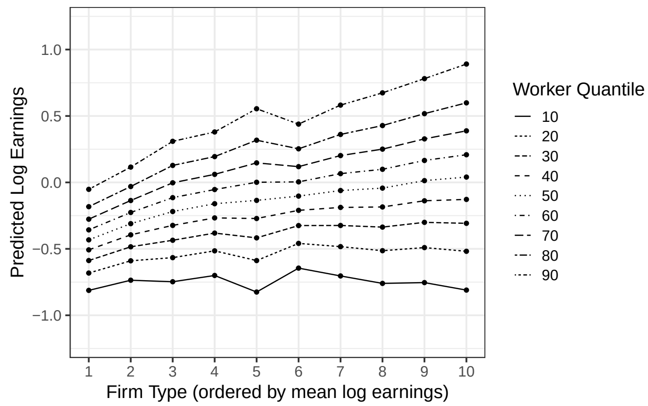


Figure A.12: Assessing Identifying Assumptions: Evidence on Firm-Worker Interactions

Notes: In this figure, we present estimates of interactions between firm and worker effects using the BLM estimator. We plot the means of log earnings for each firm type and deciles of worker heterogeneity. On the x-axis, firm types are ordered in ascending order of mean log earnings.

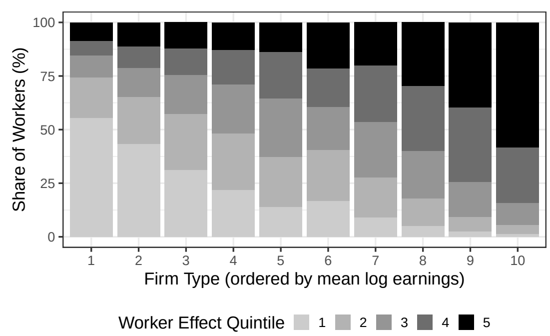


Figure A.13: Sorting of Firms and Workers

Notes: In this figure, we divide workers into 5 quintile bins and compute the share of workers in each quintile bin by firm class. We plot the firm classes in increasing order of firm effects.

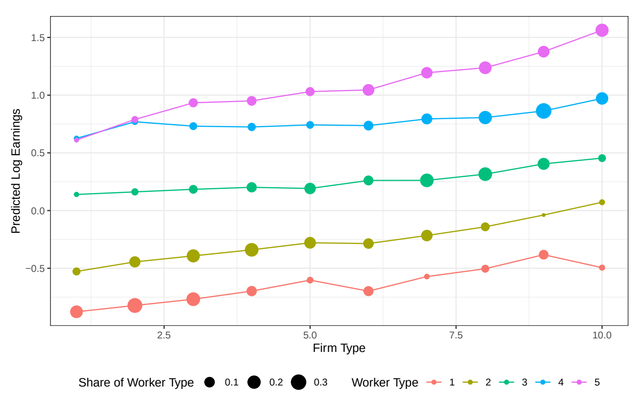
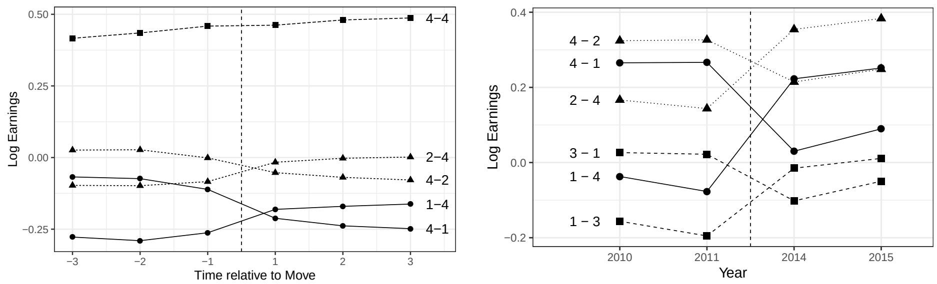


Figure A.14: Robustness to BLM Estimator with Discrete Worker Types

Notes: This figure plots predicted log earnings when using the BLM estimator with 5 discrete worker types. "Share of Worker Type" refers to the distribution of a given worker type across firm types.



(a) Baseline Sample (b) Broad Sample

Figure A.15: Assessing Identifying Assumptions: Evidence on Symmetric Changes around the Move

Notes: In this figure, we classify firms into four equally sized groups based on the mean earnings of stayers in the firm (with 1 and 4 being the group with the lowest and highest mean earnings, respectively). We then compute mean log earnings for the workers that move between these groups of firms in the years before and after the move. Note that the employer differs between event times -1 and 1, but we do not know exactly when the change in employer occurred. Thus, to avoid concerns over workers exiting and entering employment during these years, one might prefer to compare earnings in event years -2 and 2.

		Model Specification			
		(1)	(2)	(3)	(4)
Share explained by:					
i) Worker Quality	$Var(x_i)$	72.4%	70.4%	73.5%	71.6%
ii) Firm Effects	$Var(\psi_{j(i)})$	3.2%	4.3%	3.0%	4.3%
iii) Sorting	$2Cov(x_i, \psi_{j(i)})$	12.9%	13.1%	12.8%	13.1%
iv) Interactions	$Var(\varrho_{ij})$		3.0%		3.3%
	$+2Cov(x_i + \psi_{j(i)}, \varrho_{ij})$		-1.8%		-2.5%
v) Time-varying Effects	$Var(\psi_{j(i),t} - \psi_{j(i)})$			0.3%	0.3%
	$+2Cov(x_i, \psi_{j(i),t} - \psi_{j(i)})$				
Sorting Correlation:	$Cor(x_i, \psi_{j(i)})$	0.43	0.38	0.43	0.37
Variance Explained:	R^2	0.89	0.89	0.90	0.90
Specification:					
Firm-Worker Interactions		✗	✓	✗	✓
Time-varying Firm Effects		✗	✗	✓	✓

Table A.8: Comparison of BLM Specifications

Notes: This table presents the decomposition of log earnings variation into firm and worker effects using the BLM estimator for four specifications: baseline, allowing for worker effects to interact with firm effects ("Firm-Worker Interactions"), allowing for a time-varying component in the firm effects due to the pass through of value added shocks ("Time-varying Firm Effects"), and allowing for both interactions between firm and worker effects and time-varying firm effects. The analysis uses both workers who move between firms and stayers.

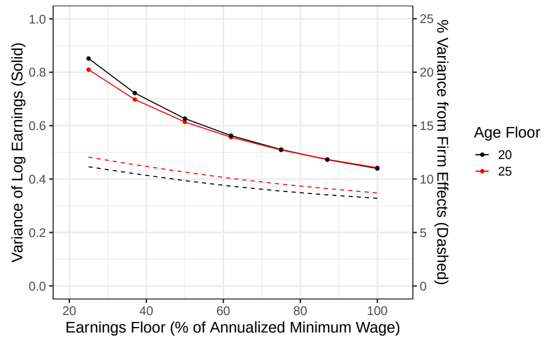


Figure A.16: Comparison of Log Earnings Variance by Earnings Floor

Notes: In this figure, we re-construct the full sample used to estimate the variance of log earnings (left y-axis) and the variance of AKM firm effects (right y-axis) when imposing different earnings floors and different age floors.

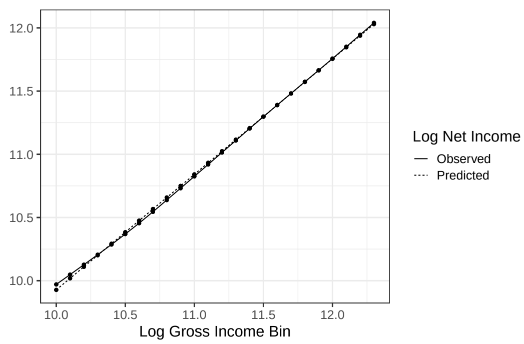


Figure A.17: Fit of the Tax Function

Notes: In this figure, we display the log net income predicted by the tax function compared to the log net income observed in the data.

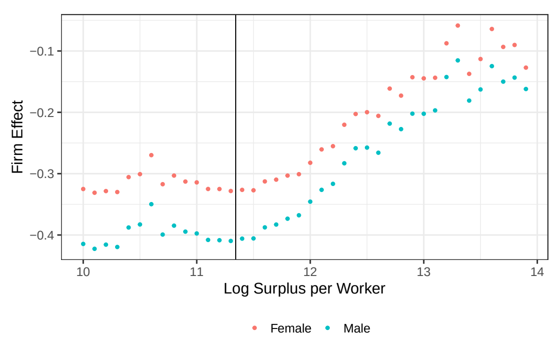


Figure A.18: Comparison of Firm Effects by Gender

Notes: This figure plots AKM firm effects for females and males against log value added per worker. On average, the difference in firm effects conditional on being male (female) is 0.0137 (0.0128) if using AKM and 0.0168 (0.0099) if using BLM-interacted. The difference in sorting conditional on being male (female) is 0.0291 (0.0300) if using AKM and 0.0276 (0.0345) if using BLM-interacted.

Set:									
Baseline Years		2001-2006			2010-2012			2010-2015	
Full Set	✓	×	×	✓	×	×	✓	×	×
Connected Set	×	✓	×	×	✓	×	×	✓	×
Leave-one-out Set	×	×	✓	×	×	✓	×	×	✓
Sample Counts (1,000):									
Unique Firms	6,717	2,954	2,009	8,870	1,241	670	7,565	2,568	1,689
(Share of Full Set)	(100%)	(44%)	(30%)	(100%)	(14%)	(8%)	(100%)	(34%)	(22%)
Unique Workers	63,146	59,748	57,027	44,182	36,826	33,031	59,621	55,464	52,484
(Share of Full Set)	(100%)	(95%)	(90%)	(100%)	(83%)	(75%)	(100%)	(93%)	(88%)
Distribution of Moves:									
Moves per Firm	2.9	6.6	9.2	0.5	3.4	5.4	2.0	5.8	8.3
Worker-weighted quantiles:									
10th Quantile	4.0	4.0	6.0	2.0	2.0	3.0	3.0	4.0	5.0
50th Quantile	72.0	73.0	82.0	23.0	25.0	33.0	56.0	58.0	67.0
90th Quantile	6,226.1	6,277.1	6,560.6	1,604.3	1,649.5	1,822.5	4,214.2	4,304.3	4,675.8
Log Earnings Distrib.:									
Variance	0.397	0.395	0.395	0.432	0.436	0.440	0.413	0.414	0.416
Between-firm Share	34%	34%	33%	39%	38%	38%	40%	40%	39%

(a) Sample Characteristics

Panel A.	**Share of Total Variation**								
Years:		**2001-2006**			**2010-2015**			**2010-2015**	
	FE	**FE-HO**	**CRE**	**FE**	**FE-HO**	**CRE**	**FE**	**FE-HO**	**CRE**
Firm Effects	12.8%	6.5%	6.4%	16.3%	4.1%	5.2%	12.2%	5.5%	6.2%
Sorting	-0.7%	10.6%	12.1%	-12.0%	11.7%	12.5%	1.1%	13.5%	15.0%
Posterior Firm Effects			7.1%						6.7%
Panel B.	**Share of Between Firm Variation**								
Years:		**2001-2006**			**2010-2015**			**2010-2015**	
	FE	**FE-HO**	**CRE**	**FE**	**FE-HO**	**CRE**	**FE**	**FE-HO**	**CRE**
Firm Effects	37.3%	19.1%	18.7%	42.8%	10.7%	13.8%	30.9%	13.9%	15.8%
Sorting	-1.9%	31.1%	35.2%	-31.3%	30.7%	32.7%	2.7%	34.1%	38.0%
Segregation	64.6%	49.7%	46.1%	88.5%	58.6%	53.5%	66.3%	51.9%	46.3%

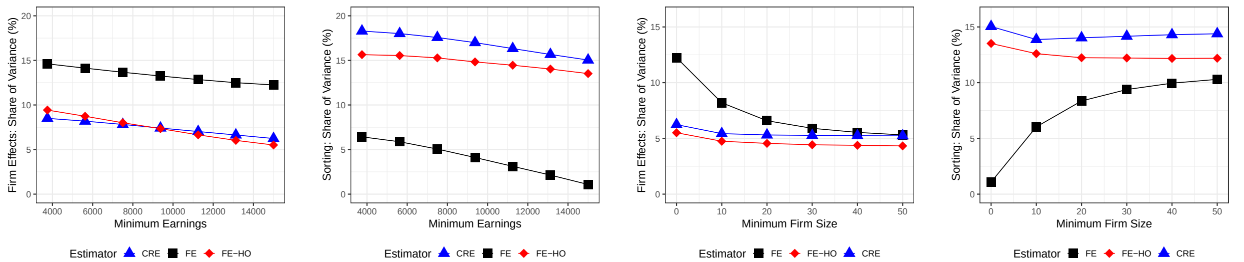
(b) Connected Set

Panel A.	**Share of Total Variation**											
Years:	**2001-2006**				**2010-2012**				**2010-2015**			
	FE	**FE-HO**	**FE-HE**	**CRE**	**FE**	**FE-HO**	**FE-HE**	**CRE**	**FE**	**FE-HO**	**FE-HE**	**CRE**
Firm Effects	10.2%	6.4%	6.7%	6.2%	10.4%	4.3%	4.5%	5.0%	9.5%	5.5%	5.8%	5.9%
Sorting	3.7%	10.4%	9.9%	11.7%	-0.8%	11.0%	10.6%	12.1%	5.9%	13.0%	12.5%	14.6%
Posterior Firm Effects				6.8%				5.2%				6.4%
Panel B.	**Share of Between Firm Variation**											
Years:	**2001-2006**				**2010-2012**				**2010-2015**			
	FE	**FE-HO**	**FE-HE**	**CRE**	**FE**	**FE-HO**	**FE-HE**	**CRE**	**FE**	**FE-HO**	**FE-HE**	**CRE**
Firm Effects	30.5%	19.3%	20.0%	18.6%	27.7%	11.3%	11.9%	13.2%	24.3%	14.2%	14.9%	15.3%
Sorting	11.2%	31.2%	29.8%	35.0%	-2.1%	29.3%	28.2%	32.2%	15.1%	33.5%	32.2%	37.6%
Segregation	58.3%	49.5%	50.3%	46.4%	74.4%	59.4%	59.9%	54.6%	60.6%	52.3%	52.9%	47.1%

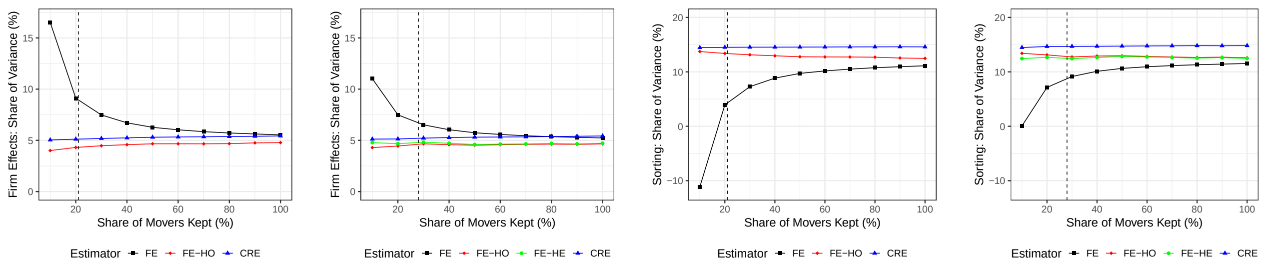
(c) Leave-one-out Set

Table A.9: Broad Sample - Total and Between Decompositions for Various Estimators

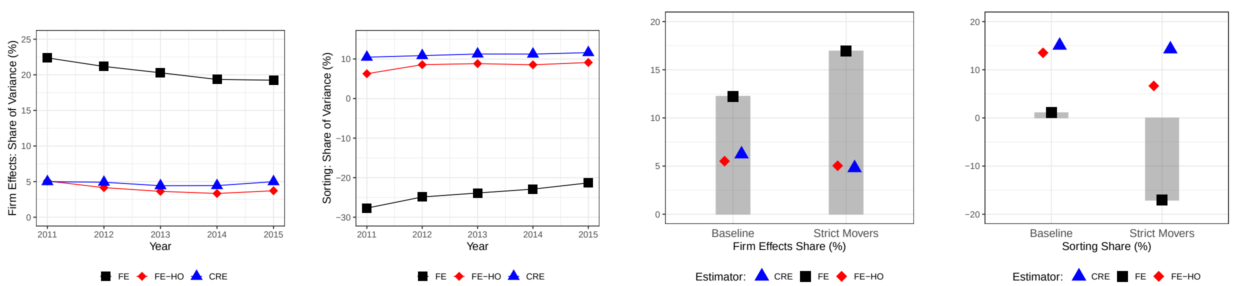
Notes: This table presents the decomposition of log earnings variation within and between firms using various estimators for three time intervals in the Broad Sample. The analysis uses both workers who move between firms and stayers.



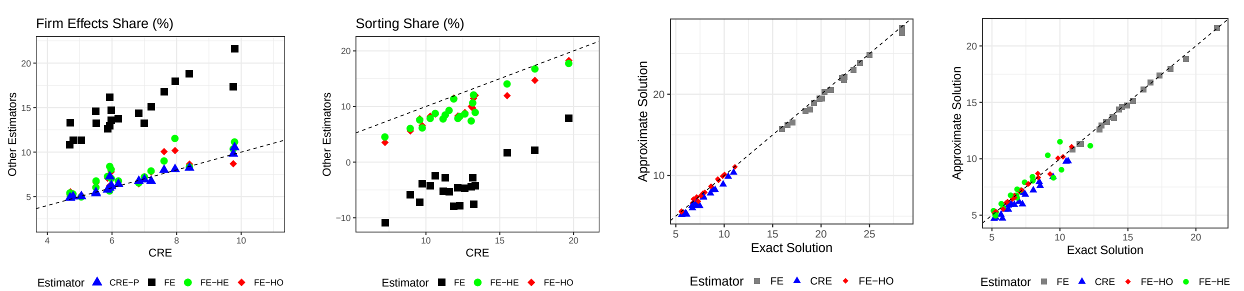
(a) Earnings Threshold: Firm Effects (b) Earnings Threshold: Sorting (c) Firm Size Threshold: Firm Effects (d) Firm Size Threshold: Sorting



(e) Limited Mobility Bias: Firm Effects (connected) (f) Limited Mobility Bias: Firm Effects (leave-one-out) (g) Limited Mobility Bias: Sorting (connected) (h) Limited Mobility Bias: Sorting (leave-one-out)



(i) Short Panels: Firm Effects (j) Short Panels: Firm Effects (k) Strict Movers: Firm Effects (l) Strict Movers: Sorting



(m) States: Firm Effects (leave-one-out) (n) States: Sorting (leave-one-out) (o) States: Approximation Methods (connected) (p) States: Approximation Methods (leave-one-out)

Figure A.19: Broad Sample - Various Robustness Checks

Notes: In this figure, we repeat several exercises shown previously, but now applied to the 2010-15 Broad Sample instead of the Baseline Sample (unless otherwise noted). See the text for details on the robustness exercises.

D Appendix: Analyses based on Exogenous Income Shocks

D.1 Identifying lottery-induced income effects

We begin by laying out the individual comparisons that allow us to estimate the effect of lottery income shocks by comparing lottery winners with earlier win years to those with later win years. Throughout, when we refer to individuals who win the lottery in the same calendar year, we will call them a cohort. To start, let Y_{it} be an outcome (such as earnings) observed for individual i in year t. Let E_i denote the year when individual i receives a lottery income shock. As a running example, consider the simple case with two groups – a group of winners winning in the year 2003 (the $E_i = 2003$ cohort) and all later winners (the $E_i > 2003$ cohorts). Suppose that for individuals in both groups, we have outcome data on year 2002 and 2003. In principle, we could calculate the average difference over time for the $E_i = 2003$ cohort and deem the average year-on-year change (e.g., $\mathbb{E}\left[Y_{i2003} - Y_{i2002}\right| E_i = 2003]$) as the on-impact effect of lottery income shocks on the outcome. However, we may worry that events coinciding with the lottery or pre-existing trends in the outcome would bias our estimate of the effect of lottery income on the outcome with such a single difference. The existence of the later-winning group helps resolve these issues to the extent that it experiences similar evolution in the outcome over time (i.e., a parallel trend). Such a cohort is useful because its outcome is also observed in both 2002 and 2003, but it only receives a lottery shock in 2004 or later, meaning that any year-on-year change in the outcome for the $E_i > 2003$ group in 2002 and 2003 will be due to the pre-existing factors but not the (future) income shock.[2] Together, these suggest a difference-in-differences strategy to recover the effect of a lottery shock for the $E_i = 2003$ cohort in 2003:

$$\mathbb{E}\left[Y_{i2003} - Y_{i2002}\right| E_i = 2003] - \mathbb{E}\left[Y_{i2003} - Y_{i2002}\right| E_i > 2003] \tag{14}$$

With the availability of data on subsequent calendar years, we could recover a set of dynamic effects for the $E_i = 2003$ in all calendar years for which there remains a not-yet-treated cohort of individuals. Taking this logic further, for each E_i with any later-treated cohorts, we could estimate a set of cohort-specific dynamic effects. In our subsequent event study estimates, we use all winning cohorts with all available later winners, holding the baseline pre-treatment year as two years pre-win. We produce these estimates using a sample that begins with the universe of recipients of a Form W-2G with recorded state lottery income between 2001 and 2016. We first restrict attention to winners who also have recorded age and sex data available from SSA

[2] In the presence of possible anticipation of future lottery win, some of the year-on-year changes in the $E_i >$ 2003 group's outcome may be due to the anticipatory response in addition to pre-existing trends. We abstract from this in our discussion, but note that one natural approach to acknowledge the potential for such anticipation would be to allow for an anticipation window for later-winning cohorts and only use their observations in the period prior to the onset of anticipation. We take this anticipation window approach when producing event study estimates for the change in asset value and consumption expenditure, both of which include a first-difference term and hence mechnically includes something akin to "anticipation."

records. We then focus on individuals who were 21 to 64 years of age (i.e., working age) at the time of receiving the Form W-2G, and whose first recorded W-2G state lottery payment between 2001 and 2016 was for $30,000 or more. We use this first state lottery payment to define the size of the shock that an individual experienced; however, all individuals in the sample receive a lottery-induced income shock at some point between 2001 and 2016. See Table A.10 for an overview of the descriptive characteristics, and Figure A.20 for a motivation of the research design and main event study estimates.

D.2 Income shocks on earnings, employment, and related outcomes

We proceed to a discussion of our main findings on the effect of lottery income shocks on employment, earnings, and related outcomes. Starting with wage earnings of the winner, we find that cohorts shocked with lottery income decrease their earnings by approximately $4000 one year after winning, with relatively little change (increase or decrease) in this impact in the subsequent 4 years. The analogous employment response is a 4 percentage point reduction in employment, with effects growing over time. Together with these responses in the labor market, individuals shocked with lottery income appear to increase their asset holdings as reflected by their increase in reported capital income (coming largely in the form of interest-bearing assets) and to increase their total expenditure (as implied by their chainge in capital income, once capitalized).[3] One of the key behavioral response margins, however, is geographic mobility. In particular, individuals do move in response to winning, but it is less clear what characterizes these moves (Figures A.21-A.22). As individuals win lottery prizes of varying sizes, we normalize the above effects by scaling them by the mean post-tax adult-equivalent lottery winnings, constructing effects per-dollar won (post-tax) in Table A.11. In order to capture two potentially important sources of heterogeneity, we explore these per-dollar effects accomodating variation in patterns over time (short run of years 1 and 2, and long run of years 3 to 5) and well as across pre-win incomes (quartiles of adjusted gross income).[4] Starting, as before, from the labor market outcomes, we see that wage earnings decline by approximately $0.02 per post-tax dollar won. However, this masks heterogeneity in earnings responses across the income distribution – per-dollar effects increase (in absolute value) from $0.01 per post-tax dollar won up to $0.03 per post-tax dollar won as we move across income quartiles. This pattern of increasing (in absolute value) per dollar effects is reversed for employment effects (which we scale by 100,000 for readability). Of this total effect we report a non-negligible share is represented by extensive margin, whether retiring or leaving the labor force (Table A.12) or moving into alternative work (Table A.13).

[3] For capitalizing capital income into asset value, we use the mean rate of return over time of 0.054, as calculated from the supplementary materials of Saez and Zucman (2016). We also explore alternative choices of capitalization rate; namely 0.07 and 0.10.

[4] As we do not observe precisely the date that an individual wins, we focus on effects starting from the first complete calendar year after the win year.

D.3 Heterogeneity in response

Marital status. Given that winners vary in terms of their marital status at the time of win, two natural questions to ask are: 1) do the earnings responses document above vary by marital status, and 2) are there marital responses to winning in the form or new marriages or family dissolutions? While married winners tend to have, if anything, smaller aggregate responses to winnings, once we consider that household resources are shared, we find that per-dollar, married winners are actually more responsive, decreasing their earnings by approximately \$0.03 per post-tax dollar won while singles reduce their earnings by approximately \$0.02 per post-tax dollar won (Table A.14). Beyond larger earnings responses, we also observe that winning tends induce new marriages (among single winners) and preserve existing marriages (among married winners). In light of these findings, we ask: among married winners, who (within the household) is the most responsive? In Table A.15, we find that the winner is nearly three times as responsive as the spouse.

Gender. A large existing literature documents differential responses of men and women to changes in wages and other determinants of labor supply and earnings. To explore if this heterogeneity also arises in terms of responses to an exogenous income shock, we explore heterogeneity in responses by gender of the winner. We report our findings in Table A.15: male winners are significantly more responsive (per dollar won) than female winners.

D.4 From per-dollar effects to per-period income effects

While lottery income shocks present a unique opportunity to study income effects, in order to make them more comparable to the kind of income shocks in our modelling approach, we need to translate the one-time wealth shock of a lottery prize into a change of per-period income. To do so, we take two approaches. In the first, we suppose that individuals live to the age of 80 and convert the one-time post-tax lottery win into a stream of fixed annuity payments (with an internal rate of return of 2.5%, approximately matching inflation-adjusted risk-free Treasury securities). In the second, we utilize observed sources of unearned income (such as dividends, interest payments, and rental and royalty income, among others) and use the capitalization approach (Saez and Zucman (2016)) to recover the winner's chosen allocation of unearned income over time. With both per-period income concepts, we form ratios as we did in the prior section, recovering per-period income effects for the set of outcomes explored before. We summarize the per-period income effects when we take the annuitization approach in Figure A.23. In particular, we can highlight that for each dollar of income from the lottery annuity, earnings fall by roughly \$0.50, and expenditure increases by \$0.60. As with the per-dollar effects, per-period income effects on earnings increase with pre-win income, whereas per-period income effects on expenditure decrease with pre-win income. When we instead use the second per-period income approach (allocated unearned income), we find that across outcomes of interest, per-period income effects are very similar. In the remaining Tables A.16-A.24, we demonstrate that the pattern of responses across the additional margins and subgroups discussed above are robust to modeling per-dollar effects as per-period income.

D.5 Tables and Figures

Table A.10: Descriptive statistics

(a) Winners Sample versus Broader Population

		Winners (Age 21-64)	**Tax Filers (Age 21-64)**
		(1)	(2)
Covariate	*Statistic*		
Wage Earnings	Mean	\$34,541	\$33,005
Employment	Prop.	0.79	0.80
Age	Mean	43.93	41.78
Female	Prop.	0.39	0.51
Married	Prop.	0.45	0.58
Homeowner	Prop.	0.45	0.49
Relative Q1 AGI Share		0.28	0.25
Relative Q2 AGI Share		0.21	0.25
Relative Q3 AGI Share		0.24	0.25
Relative Q4 AGI Share		0.27	0.25
N		90,731	154,372,671

		Treatment Group (Current Winners)	**Control Group (Not-Yet Winners)**
		(1)	(2)
Covariate	*Statistic*		
Wage Earnings	Mean	\$34,649	\$34,278
Employment	Prop.	0.80	0.80
Age	Mean	43.94	41.84
Female	Prop.	0.39	0.39
Married	Prop.	0.45	0.45
Homeowner	Prop.	0.45	0.44
Size of the Lottery Win	Mean	\$182,902	\$184,184

(b) Winners Sample versus Later Winners (Control Units)

Prize Range	**Count**	**Mean Pre-Tax Winnings** (per winner)	**Median Pre-Tax Winnings** (per winner)
	(1)	(2)	(3)
< 50K	24,800	\$36,512	\$34,700
50K to 100K	28,689	\$63,374	\$59,400
100K to 200K	17,521	\$126,020	\$119,100
200K+	19,721	\$1,405,008	\$307,200
All Winners	90,731	\$359,743	\$67,800

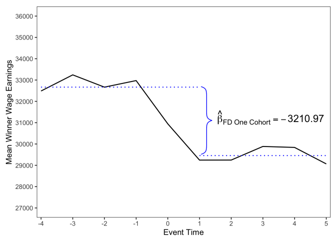
(c) Distribution of Prize Sizes

		Winner	**Spouse**
		(1)	(2)
Covariate	*Statistic*		
Wage Earnings	Mean	\$42,465	\$33,037
Employment	Prop.	0.80	0.73
Age	Prop.	47.07	46.65
Female	Prop.	0.36	0.64
Primary Earner	Prop.	0.62	0.38
Older Member	Prop.	0.50	0.50
Same Age	Prop.	0.11	0.11
		$p < 0.001$	

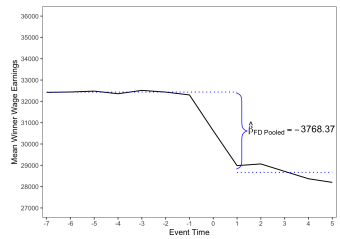
(d) Winner versus Spouse

Notes: These tables presents a summary of the descriptive statistics in our sample of working-age winners. All monetary values are reported in 2016 U.S. dollars, using the Consumer Price Index to adjust for inflation. The control gorup consists of later winners. Medians rounded to nearest hundred.

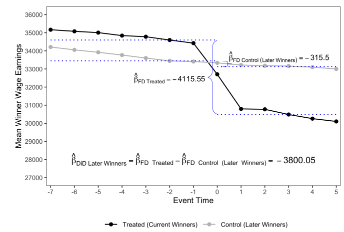
Figure A.20: Motivation for Research Design and Responses



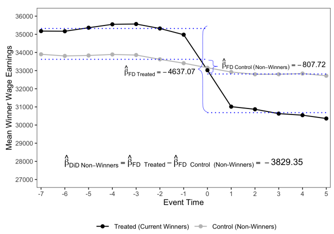
(a) First-Difference (Single Cohort)



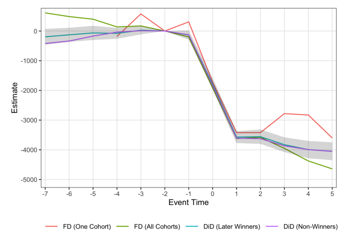
(b) First-Difference (Pooled Across Cohorts)



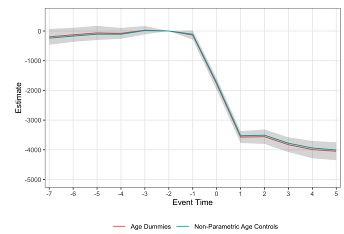
(c) Difference-in-Differences (Using Later Winners as Controls)



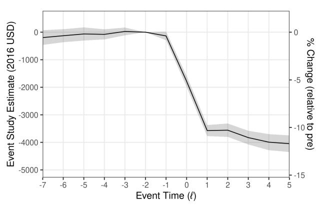
(d) Difference-in-Differences (Using Non-Winners as Controls)



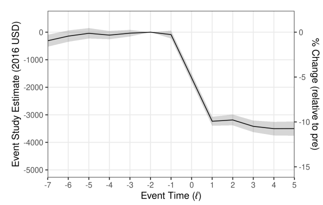
(e) Comparison of Estimators



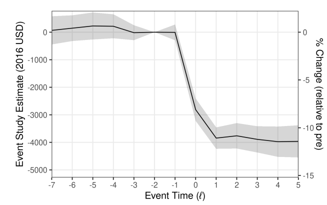
(f) Comparison of Approaches to Control for Life-Cycle Effects



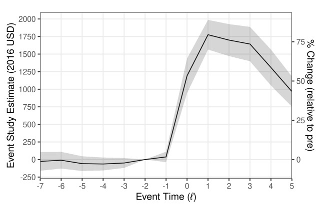
(g) Lottery-Induced Income Effect on Wage Earnings



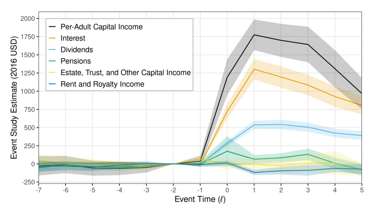
(h) Lottery-Induced Income Effect on Per-Adult Wage Earnings



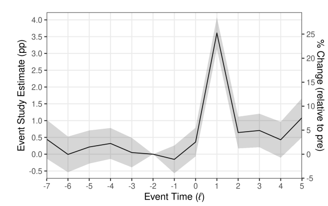
(i) Lottery-Induced Income Effect on Per-Adult Total Earnings (Wage + Self-Employment Income)



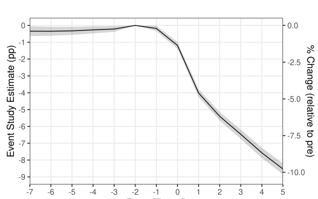
(j) Lottery-Induced Income Effect on Per-Adult Capital Income



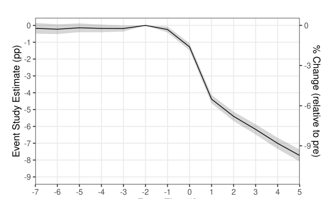
(k) Lottery-Induced Income Effect on Per-Adult Capital Income



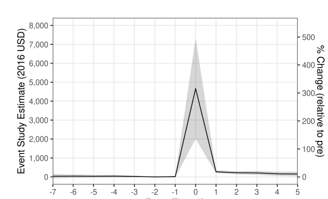
(l) Lottery-Induced Income Effect on Geographic Mobility



(m) Lottery-Induced Income Effect on Total Employment



(n) Lottery-Induced Income Effect on Winner Employment

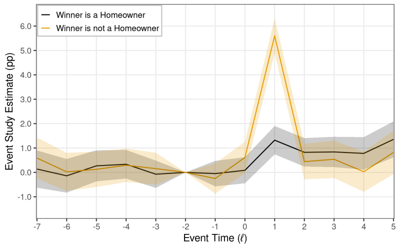


(o) Lottery-Induced Income Effect on Charitable Contributions

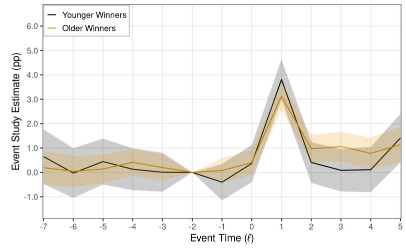
Notes: These figure provides a comparison of various estimators for the effect of winning the lottery on earnings. In addition, we present event study estimates of the effect of lottery-induced shocks on wage earnings and employment, as well as several outcomes which might have behavioral responses linked to earnings / employment responses. 90% confidence intervals are displayed for event studies, clustering on individual winners.

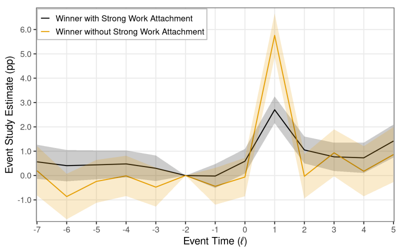
Figure A.21: Geographic Mobility and Decompositions by Subgroups

(a) Homeowners and renters

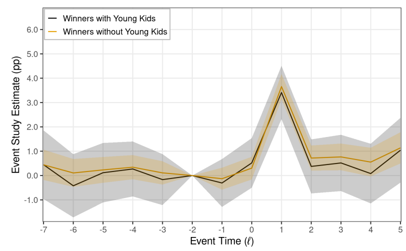


(b) Young and old winners

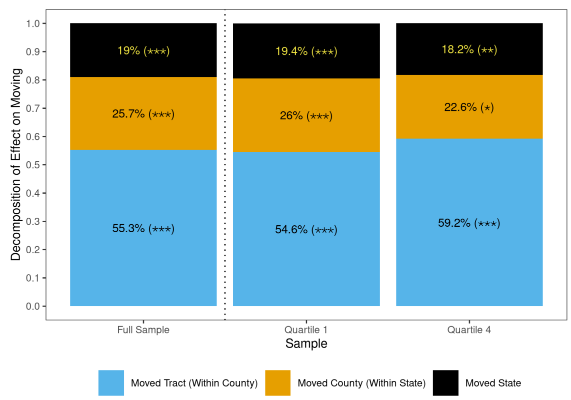




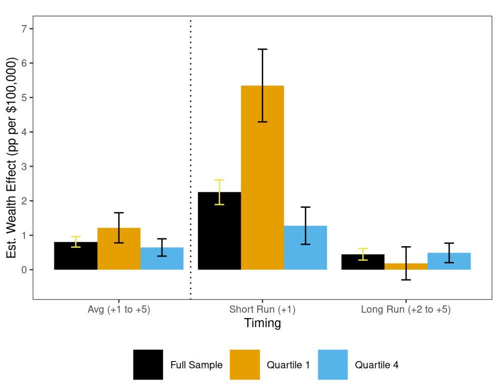
(c) Work attachment

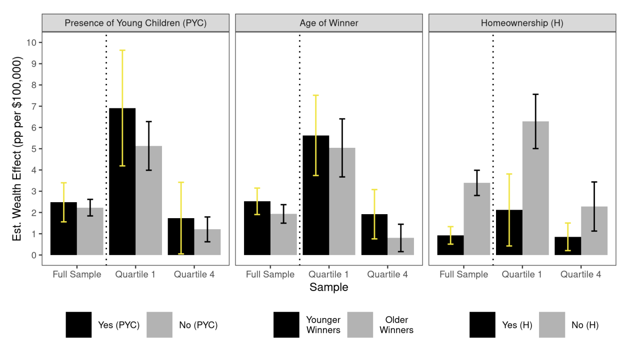


(d) Parents with young kids



(e) Decomposition by Distance Moved



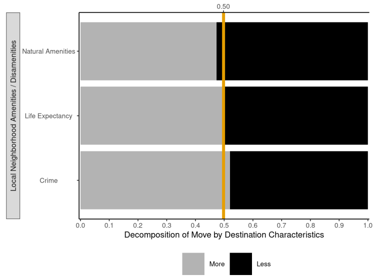


(f) Wealth Effects

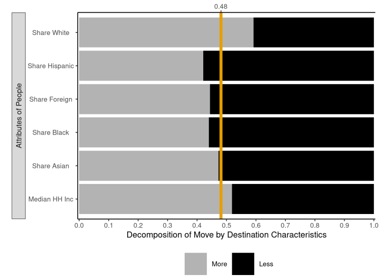
(g) Wealth Effects by Subgroup

Notes: This figure presents estimates of the impact of winning on the propensity to move across Census tracts for several subgroups, as well as decompositions of where they move. All outcomes are deteremined using Census tracts. 90 percent confidence intervals are displayed for event studies, clustering on winner.

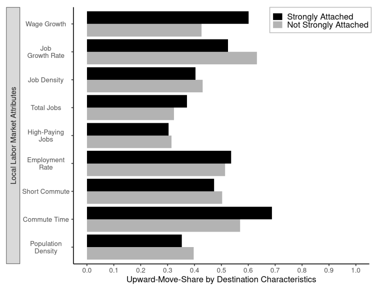
Figure A.22: Geographic Mobility and Decompositions by Subgroups



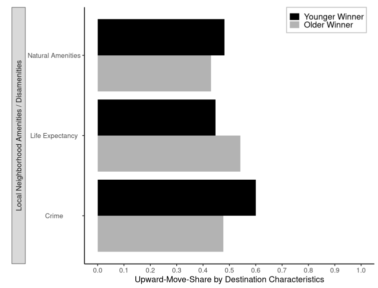
(a) Amenities and Disamenities



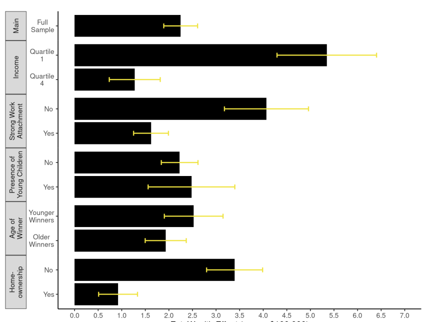
(b) Attributes of Neighbours



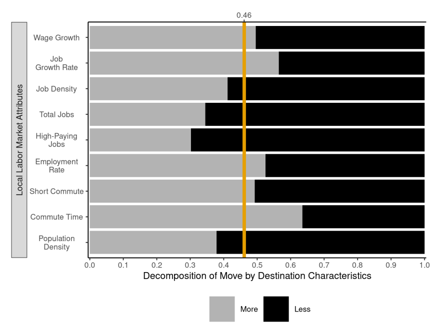
(c) Local Labor Market Attributes for Winners by Labor Market Attachment



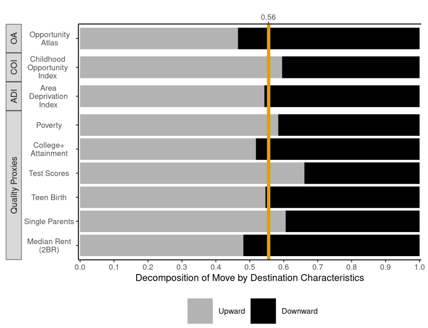
(d) Amenities and Disamenities by Age of Winner



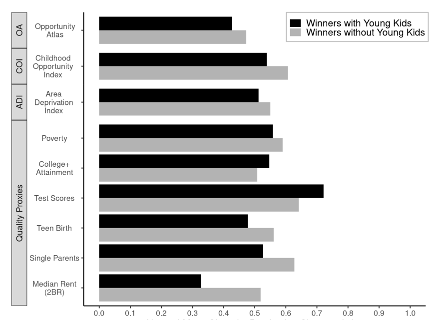
(e) Winner Characteristics



(f) Destination Characteristics: Local Labor Market



(g) Destination Characteristics: Neighborhood Quality



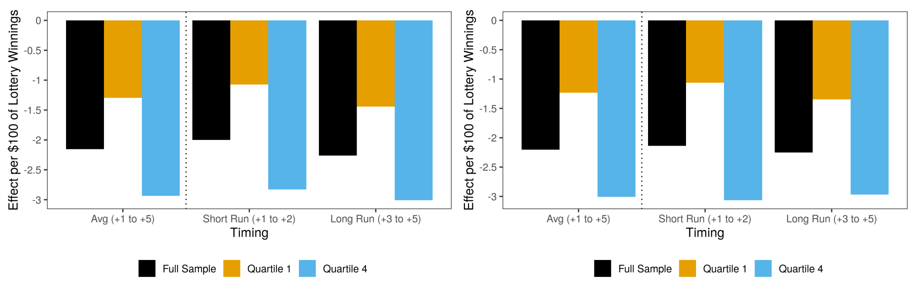
(h) Neighborhood Characteristics for Winners with Young Children

Notes: This figure presents estimates of the impact of winning on the propensity to move across Census tracts for several subgroups, as well as decompositions of where they move. All outcomes are deteremined using Census tracts.

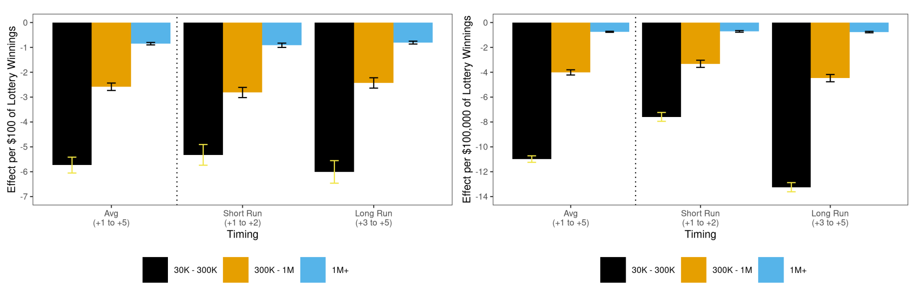
Table A.11: Wealth effects across outcomes

	Sample				
Outcome	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 2 Pre-Win Income (3)	Quartile 3 Pre-Win Income (4)	Quartile 4 Pre-Win Income (5)
Winner Wage Earnings (per $100)	-2.2856 (0.0571)	-1.4003 (0.0628)	-2.2948 (0.0861)	-2.6196 (0.0935)	-3.0596 (0.1541)
Per-Adult Wage Earnings (per $100)	-2.0245 (0.0492)	-1.2514 (0.0589)	-2.0422 (0.0764)	-2.2590 (0.0813)	-2.7035 (0.1311)
Per-Adult Total Labor Earnings (per $100)	-2.3394 (0.0657)	-1.3339 (0.1051)	-2.2720 (0.0867)	-2.6450 (0.0996)	-3.1298 (0.1820)
Per-Adult Capital Income (per $100)	0.8738 (0.0406)	0.5784 (0.0540)	0.7626 (0.0784)	0.9658 (0.0709)	0.9265 (0.0974)
Winner Employment (per $100,000)	-0.0368 (0.0008)	-0.0517 (0.0021)	-0.0444 (0.0019)	-0.0350 (0.0013)	-0.0231 (0.0010)
Total Employment (per $100,000)	-0.0385 (0.0008)	-0.0639 (0.0025)	-0.0447 (0.0018)	-0.0322 (0.0012)	-0.0218 (0.0008)

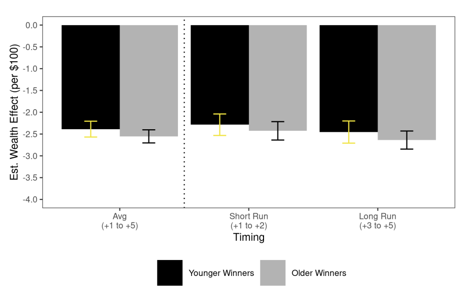
(a) Wealth effects across time and pre-win income (wage earnings and per-adult labor earnings)



(b) Wealth effects by prize size over time time and pre-win income (per-adult labor earnings and total employment)



(c) Wealth effects by age of winner (per-adult labor earnings)



Notes: These tables and figures presents estimates of the mean effect per dollar of lottery winnings (i.e., wealth effect) on six earnings and employment outcomes. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner. To ease interpretability, we scale earnings and capital income responses by $100. In the case of employment responses, we scale each estimate by $100,000. In addition, in the figures, we highlight differences in estimates over time, by income quartile, and by age of winner.

Table A.12: Extensive margin responses (level and share of earnings response)

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
		Take-Up of Retirement Benefits		
Claiming OASI Benefits	*Estimate*	0.0114	0.0224	0.0077
	Standard Error	(0.0041)	(0.0122)	(0.0077)
	Counterfactual Mean	0.77	0.74	0.73
	%Δ	1.5	3.0	1.1
		Labor Market Exit		
One-Year Exit	*Estimate*	0.0489	0.0361	0.0392
	Standard Error	(0.0053)	(0.0095)	(0.0103)
	Counterfactual Mean	0.43	0.63	0.33
	%Δ	11.3	5.8	11.7
Two-Year Exit	*Estimate*	0.0536	0.0477	0.0457
	Standard Error	(0.0058)	(0.0106)	(0.0111)
	Counterfactual Mean	0.40	0.58	0.30
	%Δ	13.4	8.3	15.0
Five-Year Exit	*Estimate*	0.0490	0.0599	0.0195
	Standard Error	(0.0098)	(0.0265)	(0.0170)
	Counterfactual Mean	0.35	0.48	0.31
	%Δ	14.0	12.6	6.4

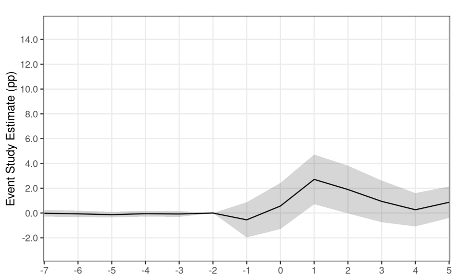
(a) Extensive margin share - Full Sample

	Sample				
Outcome	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 2 Pre-Win Income (3)	Quartile 3 Pre-Win Income (4)	Quartile 4 Pre-Win Income (5)
Winner Wage Earnings	0.60	0.65	0.50	0.52	0.50
Per-Adult Total Labor Earnings	0.54	0.59	0.44	0.41	0.41

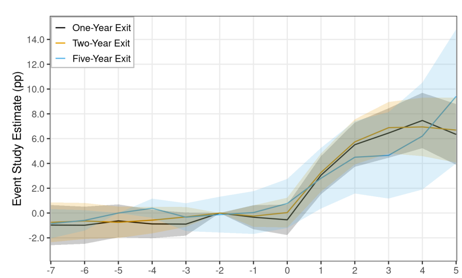
(b) Extensive margina share - Across Prize Size

	Sample				
Outcome	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 2 Pre-Win Income (3)	Quartile 3 Pre-Win Income (4)	Quartile 4 Pre-Win Income (5)
Winner Wage Earnings	0.62	0.66	0.50	0.53	0.49
Per-Adult Total Labor Earnings	0.64	0.77	0.51	0.49	0.48

(c) Claiming social security benefits



(d) Labor market exit



Notes: Thess tables and figures present levels and share reponses along the extensive margin. On top, we report estimates of the mean effect per dollar of lottery winnings (i.e., wealth effect) on take up of retirement benefits and labor market exit for winners aged 62-64. The dependent variables are binary indicators for the receipt of OASI benefits and labor force exit respectively. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner. To ease interpretability, each estimate is scaled by $100,000. Then, the two bottom tables present the share of the response due to extensive margin in aggregate and across prize sizes. 90 percent confidence intervals are displayed for event studies, clustering on winner.

Table A.13: Effects of wealth on labor market transitions

(a) Wealth effects by labor market attachment

		Sample				
Outcome	Labor Force Attachment	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 2 Pre-Win Income (3)	Quartile 3 Pre-Win Income (4)	Quartile 4 Pre-Win Income (5)
Winner Wage Earnings	Baseline	-2.2856 (0.0571)	-1.4003 (0.0628)	-2.2948 (0.0861)	-2.6196 (0.0935)	-3.0596 (0.1541)
	Employed Pre-Win	-2.6437 (0.0661)	-1.7358 (0.0968)	-2.6309 (0.0903)	-2.8243 (0.0995)	-3.3511 (0.1580)
	FT Employed Pre-Win	-2.9566 (0.0792)	-2.6015 (0.3019)	-2.9885 (0.0987)	-2.9500 (0.1046)	-3.4603 (0.1621)
Per-Adult Wage Earnings	Baseline	-2.0245 (0.0492)	-1.2514 (0.0589)	-2.0422 (0.0764)	-2.2590 (0.0813)	-2.7035 (0.1311)
	Employed Pre-Win	-2.2414 (0.0546)	-1.4446 (0.0896)	-2.2479 (0.0781)	-2.3421 (0.0847)	-2.8994 (0.1269)
	FT Employed Pre-Win	-2.4729 (0.0644)	-1.9109 (0.2741)	-2.5386 (0.0833)	-2.4170 (0.0887)	-2.9483 (0.1283)
Per-Adult Total Labor Earnings	Baseline	-2.3394 (0.0657)	-1.3339 (0.1051)	-2.2720 (0.0867)	-2.6450 (0.0996)	-3.1298 (0.1820)
	Employed Pre-Win	-2.5170 (0.0972)	-1.5688 (0.1764)	-2.2965 (0.1185)	-2.5543 (0.1320)	-3.4340 (0.2471)
	FT Employed Pre-Win	-2.7651 (0.1140)	-2.5804 (0.5710)	-2.6287 (0.1365)	-2.6342 (0.1366)	-3.4363 (0.2499)

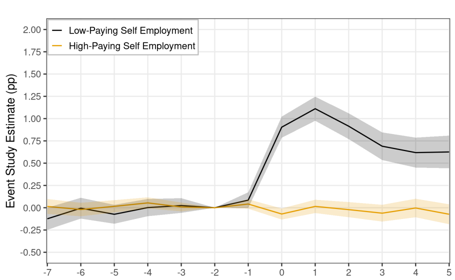
(b) Entrepreneurship and self-employment

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
Transition to Low-Paying SE	*Estimate*	0.0047	0.0037	0.0053
	Standard Error	(0.0003)	(0.0010)	(0.0004)
	Counterfactual Mean	0.03	0.05	0.03
	%Δ	13.7	6.8	19.4
Transition to High-Paying SE	*Estimate*	-0.0002	-0.0005	-0.0005
	Standard Error	(0.0002)	(0.0005)	(0.0003)
	Counterfactual Mean	0.01	0.01	0.02
	%Δ	-1.3	-3.2	-3.0

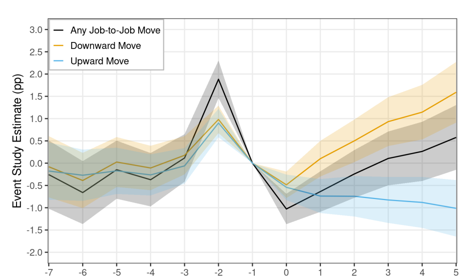
(c) Job-to-job mobility

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
Any Job-to-Job Move	*Estimate*	0.0004	0.0033	-0.0043
	Standard Error	(0.0018)	(0.0052)	(0.0030)
	Counterfactual Mean	0.46	0.67	0.36
	%Δ	0.1	0.5	-1.2
Downward Move	*Estimate*	0.0096	0.0123	0.0011
	Standard Error	(0.0017)	(0.0053)	(0.0026)
	Counterfactual Mean	0.23	0.31	0.19
	%Δ	4.3	4.0	0.6
Upward Move	*Estimate*	-0.0092	-0.0090	-0.0054
	Standard Error	(0.0016)	(0.0052)	(0.0023)
	Counterfactual Mean	0.24	0.36	0.17
	%Δ	-3.9	-2.5	-3.2

(d) Transition into self-employment



(e) Job-to-job transitions



Notes: These tables and figures presents estimates of the overall effect and mean effect per dollar of lottery winnings (i.e., wealth effect) on those strongly employed prior to winning. The responses of interest are earnings responses, as well as transitions to self employment and to a new job. The estimation sample is restricted to winners and not-yet winners in paid-employment at event time $w-2$ for most of the analysis. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner. To ease interpretability, we scale non-earnings estimates by $100,000.

Table A.14: Effects of wealth on marriages and divorce and heterogeneity by marital status

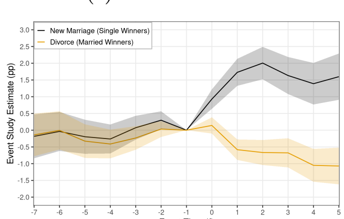
(a) Marital status

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
New Marriage	*Estimate*	0.0077	0.0167	0.0006
	Standard Error	(0.0008)	(0.0024)	(0.0013)
	Counterfactual Mean	0.14	0.14	0.15
	%Δ	5.5	11.7	0.4
Divorce	*Estimate*	-0.0067	-0.0146	-0.0058
	Standard Error	(0.0010)	(0.0041)	(0.0015)
	Counterfactual Mean	0.11	0.17	0.09
	%Δ	-5.9	-8.6	-6.3

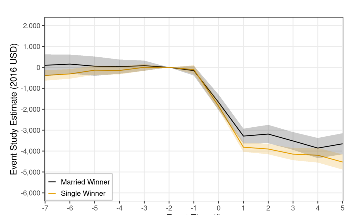
(b) Heterogeneity by marital status

Group	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
		Winner Wage Earnings (per $100)		
Married Winner	*Estimate*	-2.9677	-2.6210	-3.1574
	Standard Error	(0.1194)	(0.2250)	(0.3004)
	Counterfactual Mean	38275.25	11892.97	62152.53
	%Δ	-0.0078	-0.0220	-0.0051
Single Winner	*Estimate*	-2.0146	-1.2763	-3.0077
	Standard Error	(0.0603)	(0.0634)	(0.1749)
	Counterfactual Mean	29816.89	13198.26	57815.81
	%Δ	-0.0068	-0.0097	-0.0052
		Per-Adult Total Labor Earnings (per $100)		
Married Winner	*Estimate*	-2.8043	-2.3592	-3.2473
	Standard Error	(0.1400)	(0.5071)	(0.3593)
	Counterfactual Mean	38033.72	12365.40	63576.97
	%Δ	-0.0074	-0.0191	-0.0051
Single Winner	*Estimate*	-2.2221	-1.3374	-3.1445
	Standard Error	(0.0693)	(0.0961)	(0.1907)
	Counterfactual Mean	30994.56	14686.77	58302.67
	%Δ	-0.0072	-0.0091	-0.0054

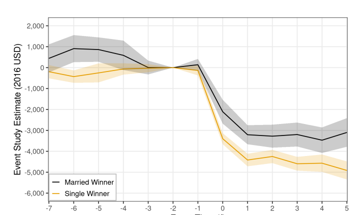
(c) Marital status



(d) Wage earnings of single and married winners



(e) Total per-adult labor income of single and married winners



Notes: These tables and figures present estimates of the overall effects and mean effect per dollar of lottery winnings (i.e., wealth effect) on the propensity to enter or leave marriage, as well as earnings effects by marital status. The estimation sample is restricted to winners and not-yet winners that are tax filers when studying entry or exit from marriage. When we study the effect on new marriages (divorce), we further restrict the sample to individuals that were not married (married) in $w-2$. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner. To ease interpretability, each estimate is scaled by $100,000.

Table A.15: Effects of wealth on earnings by winner gender and winner versus spouse

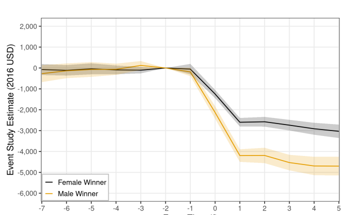
(a) Heterogeneity by gender of winner

Group	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
		Winner Wage Earnings (per \$100)		
Male Winner	*Estimate*	-2.5714	-1.6983	-3.1973
	Standard Error	(0.0826)	(0.0859)	(0.2081)
	Counterfactual Mean	38235.23	14288.29	66641.39
	%Δ	-0.0067	-0.0119	-0.0048
Female Winner	*Estimate*	-1.8539	-1.0093	-3.2610
	Standard Error	(0.0651)	(0.0785)	(0.2026)
	Counterfactual Mean	26576.05	10946.22	47967.19
	%Δ	-0.0070	-0.0092	-0.0068
		Per-Adult Total Labor Earnings (per \$100)		
Male Winner	*Estimate*	-2.7046	-1.7517	-3.5947
	Standard Error	(0.0852)	(0.1267)	(0.2065)
	Counterfactual Mean	36131.60	15015.03	62948.19
	%Δ	-0.0075	-0.0117	-0.0057
Female Winner	*Estimate*	-1.7193	-0.8651	-2.6204
	Standard Error	(0.1041)	(0.1926)	(0.3929)
	Counterfactual Mean	31112.13	12875.22	58251.54
	%Δ	-0.0055	-0.0067	-0.0045

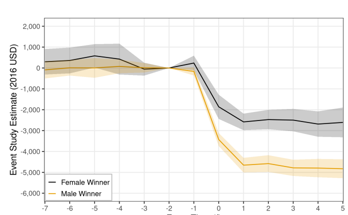
(b) Winner versus spouse responses

Group	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
Winner	*Estimate*	-2.9677	-2.6210	-3.1574
	Standard Error	(0.1194)	(0.2250)	(0.3004)
	Counterfactual Mean	38275.25	11892.97	62152.53
	%Δ	-0.0078	-0.0220	-0.0051
Spouse	*Estimate*	-1.0505	-0.3263	-1.7283
	Standard Error	(0.1032)	(0.2175)	(0.3223)
	Counterfactual Mean	27141.85	6727.89	46890.46
	%Δ	-0.0039	-0.0048	-0.0037

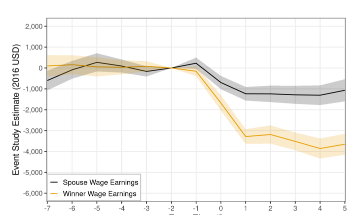
(c) Wage earnings by winner gender



(d) Total per-adult labor income by winner gender

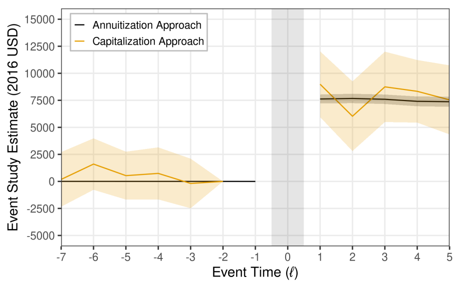


(e) Wage earnings by winner verus spouse

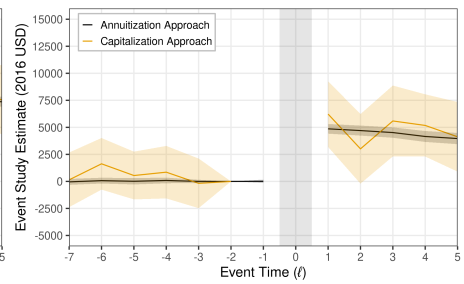


Notes: These tables and figures present estimates of the overall effects and mean effect per dollar of lottery winnings (i.e., wealth effect) on earnings separately by gender of the winner. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner.

Figure A.23: Comparing across methods to measure unearned income



(a) Unearned Income



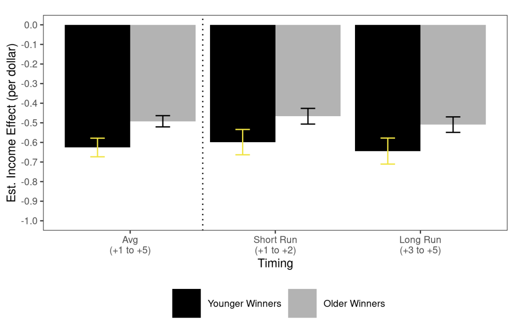
(b) Implied Consumption Expenditure

	Sample				
Outcome	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 2 Pre-Win Income (3)	Quartile 3 Pre-Win Income (4)	Quartile 4 Pre-Win Income (5)
Per-Adult Total Labor Earnings	-0.4934 (0.0191)	-0.2857 (0.0266)	-0.4838 (0.0246)	-0.5634 (0.0287)	-0.6470 (0.0542)
Per-Adult Labor Earnings Taxes	-0.1001 (0.0089)	-0.0368 (0.0127)	-0.0655 (0.0138)	-0.1198 (0.0100)	-0.1658 (0.0260)
Implied Consumption Expenditure	0.6067 (0.0272)	0.7511 (0.0501)	0.5818 (0.0454)	0.5564 (0.0424)	0.5188 (0.0644)

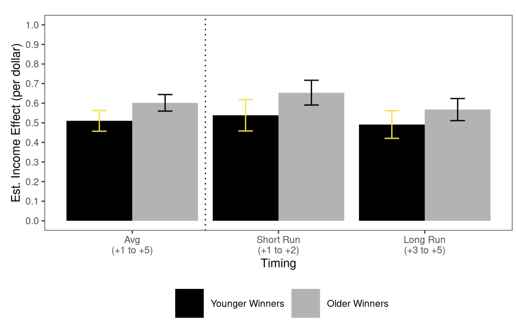
(c) IV estimates of the effect of exogenous change in unearned income

	Sample				
Outcome	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 2 Pre-Win Income (3)	Quartile 3 Pre-Win Income (4)	Quartile 4 Pre-Win Income (5)
Per-Adult Total Labor Earnings	-0.5182 (0.0815)	-0.2925 (0.0669)	-0.4278 (0.0874)	-0.5283 (0.1041)	-0.7385 (0.3306)
Per-Adult Labor Earnings Taxes	-0.1077 (0.0201)	-0.0404 (0.0221)	-0.0609 (0.0244)	-0.1143 (0.0277)	-0.1941 (0.0892)
Implied Consumption Expenditure	0.5882 (0.1689)	0.7463 (0.2294)	0.6318 (0.2280)	0.5858 (0.2159)	0.4533 (0.4781)

(d) IV estimates of the effect of exogenous change in unearned income (capitalization approach)



(f) Per-Adult Total Labor Earnings



(g) Consumption Expenditure

(g) Effects of exogenous change in unearned income by age of winner

Notes: Thes figure presents estimates of the impact of winning on unearned income and implied consumption expenditure. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner.

Table A.16: Effects of unearned income on job mobility

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
Any Job-to-Job Move	*Estimate*	0.0004	0.0033	-0.0043
	Standard Error	(0.0018)	(0.0052)	(0.0030)
	Counterfactual Mean	0.46	0.67	0.36
	%Δ	0.1	0.5	-1.2
Downward Move	*Estimate*	0.0096	0.0123	0.0011
	Standard Error	(0.0017)	(0.0053)	(0.0026)
	Counterfactual Mean	0.23	0.31	0.19
	%Δ	4.3	4.0	0.6
Upward Move	*Estimate*	-0.0092	-0.0090	-0.0054
	Standard Error	(0.0016)	(0.0052)	(0.0023)
	Counterfactual Mean	0.24	0.36	0.17
	%Δ	-3.9	-2.5	-3.2

Notes: This table presents estimates of the mean effect per dollar of unearned income on the frequency and direction of job-to-job moves. The estimation sample is restricted to winners and not-yet winners in paid-employment pre- and post-win. Outcomes are defined as binary and equal to 1 if the firm is either different from, or higher or lower ranked than the firm prior to winning the lottery. Firms are ranked by the mean wage paid to its employees. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner. To ease interpretability, each estimate is scaled by $10,000.

Table A.17: Effects of unearned income on earnings by gender of the winner

Group	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
		Winner Wage Earnings (per dollar)		
Male Winner	*Estimate*	-0.5733	-0.3934	-0.6839
	Standard Error	(0.0186)	(0.0197)	(0.0442)
	Counterfactual Mean	38235.23	14288.29	66641.39
	%Δ	-0.0015	-0.0028	-0.0010
Female Winner	*Estimate*	-0.4113	-0.2277	-0.6936
	Standard Error	(0.0144)	(0.0179)	(0.0436)
	Counterfactual Mean	26576.05	10946.22	47967.19
	%Δ	-0.0015	-0.0021	-0.0014
		Per-Adult Total Labor Earnings (per dollar)		
Male Winner	*Estimate*	-0.6031	-0.4047	-0.7683
	Standard Error	(0.0190)	(0.0290)	(0.0437)
	Counterfactual Mean	36131.60	15015.03	62948.19
	%Δ	-0.0017	-0.0027	-0.0012
Female Winner	*Estimate*	-0.3817	-0.1905	-0.5588
	Standard Error	(0.0234)	(0.0442)	(0.0856)
	Counterfactual Mean	31112.13	12875.22	58251.54
	%Δ	-0.0012	-0.0015	-0.0010

Notes: This table presents estimates of the mean effect per dollar of unearned income on earnings separately by gender of the winner. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner.

Table A.18: Effects of unearned income on earnings by marital status of the winner

Group	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
		Winner Wage Earnings (per dollar)		
Married Winner	*Estimate*	-0.6376	-0.5456	-0.6728
	Standard Error	(0.0254)	(0.0471)	(0.0633)
	Counterfactual Mean	38275.25	11892.97	62152.53
	%Δ	-0.0017	-0.0046	-0.0011
Single Winner	*Estimate*	-0.4579	-0.2982	-0.6507
	Standard Error	(0.0138)	(0.0147)	(0.0375)
	Counterfactual Mean	29816.89	13198.26	57815.81
	%Δ	-0.0015	-0.0023	-0.0011
		Per-Adult Total Labor Earnings (per dollar)		
Married Winner	*Estimate*	-0.6009	-0.4818	-0.6892
	Standard Error	(0.0300)	(0.1096)	(0.0759)
	Counterfactual Mean	38033.72	12365.40	63576.97
	%Δ	-0.0016	-0.0039	-0.0011
Single Winner	*Estimate*	-0.5051	-0.3117	-0.6836
	Standard Error	(0.0158)	(0.0222)	(0.0408)
	Counterfactual Mean	30994.56	14686.77	58302.67
	%Δ	-0.0016	-0.0021	-0.0012

Notes: This table presents estimates of the mean effect per dollar of unearned income on earnings separately by marital status of the winner. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner.

Table A.19: Effects of unearned income on wage earnings of winners and their spouse

Group	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
Winner	*Estimate*	-0.6376	-0.5456	-0.6728
	Standard Error	(0.0254)	(0.0471)	(0.0633)
	Counterfactual Mean	38275.25	11892.97	62152.53
	%Δ	-0.0017	-0.0046	-0.0011
Spouse	*Estimate*	-0.2249	-0.0706	-0.3706
	Standard Error	(0.0221)	(0.0452)	(0.0668)
	Counterfactual Mean	27141.85	6727.89	46890.46
	%Δ	-0.0008	-0.0010	-0.0008

Notes: This table presents estimates of the mean effect per dollar of unearned income on wage earnings for winners and non-winning spouses. The estimation sample is restricted to married couples. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner.

Table A.20: Effects of unearned income on earnings by identity of winner

Conditioning	Value	**Gender of Winner** Male Winner (1)	Female Winner (2)	**Relative Wage Earnings of Winner** Primary Earner (3)	Secondary Earner (4)	**Relative Age of Winner** Older (5)	Younger (6)
Two-Earner Household	Estimate	-0.7829	-0.7882	-0.7437	-0.8698	-0.8690	-0.6709
	Standard Error	(0.0409)	(0.0557)	(0.0425)	(0.0519)	(0.0443)	(0.0484)

Notes: This table presents estimates of the mean effect per dollar of unearned income on total per-adult labor earnings by identity of the winner. The estimation sample is restricted to two-earner households. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner.

Table A.21: Effects of unearned income and wealth on charitable giving

(a) Effect of unearned income

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
Reported Charitable Donations	*Estimate*	0.0196	0.0164	0.0306
	Standard Error	(0.0031)	(0.0068)	(0.0080)
	Counterfactual Mean	1546.09	1213.07	2355.03
	%Δ	< 0.1	< 0.1	< 0.1

(b) Effect of wealth

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
Reported Charitable Donations	*Estimate*	0.0908	0.0799	0.1454
	Standard Error	(0.0146)	(0.0333)	(0.0378)
	Counterfactual Mean	1546.09	1213.07	2355.03
	%Δ	< 0.1	< 0.1	< 0.1

Notes: This table presents estimates of the mean effect per dollar of wealth or unearned income on charitable giving. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner.

Table A.22: Effects of unearned income on take-up of retirement benefits and labor market exit

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
		Take-Up of Retirement Benefits		
Claiming OASI Benefits	*Estimate*	0.0169	0.0333	0.0114
	Standard Error	(0.0061)	(0.0181)	(0.0115)
	Counterfactual Mean	0.77	0.74	0.73
	%Δ	2.2	4.5	1.6
		Labor Market Exit		
One-Year Exit	*Estimate*	0.0720	0.0532	0.0581
	Standard Error	(0.0078)	(0.0140)	(0.0152)
	Counterfactual Mean	0.43	0.63	0.33
	%Δ	16.7	8.5	17.4
Two-Year Exit	*Estimate*	0.0790	0.0703	0.0678
	Standard Error	(0.0086)	(0.0156)	(0.0164)
	Counterfactual Mean	0.40	0.58	0.30
	%Δ	19.7	12.2	22.3
Five-Year Exit	*Estimate*	0.0721	0.0879	0.0290
	Standard Error	(0.0144)	(0.0390)	(0.0251)
	Counterfactual Mean	0.35	0.48	0.31
	%Δ	20.6	18.5	9.5

Notes: This table presents estimates of the mean effect per dollar of unearned income on take up of retirement benefits and labor market exit for winners aged 62-64. The dependent variables are binary indicators for the receipt of OASI benefits and labor force exit respectively. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner. To ease interpretability, each estimate is scaled by $100,000.

Table A.23: Effects of unearned income on entrepreneurship and self-employment

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
Transition to Low-Paying SE	*Estimate*	0.0107	0.0092	0.0115
	Standard Error	(0.0006)	(0.0025)	(0.0009)
	Counterfactual Mean	0.03	0.05	0.03
	%Δ	31.1	17.1	42.1
Transition to High-Paying SE	*Estimate*	-0.0004	-0.0009	-0.0011
	Standard Error	(0.0004)	(0.0012)	(0.0006)
	Counterfactual Mean	0.01	0.01	0.02
	%Δ	-2.9	-6.3	-6.3

Notes: This table presents estimates of the mean effect per dollar of unearned income on the propensity to start a business associated with annual profits of $15,000 or less (low-paying SE), or a business with profits of more than $15,000 (high-paying SE). The estimation sample is restricted to winners and not-yet winners in paid-employment at event time $w-2$.We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner. To ease interpretability, each estimate is scaled by $100,000.

Table A.24: Effects of unearned income on marriages and divorce

Outcome	Value	Full Sample (1)	Quartile 1 Pre-Win Income (2)	Quartile 4 Pre-Win Income (3)
New Marriage	*Estimate*	0.0174	0.0388	0.0009
	Standard Error	(0.0018)	(0.0056)	(0.0029)
	Counterfactual Mean	0.14	0.14	0.15
	%Δ	12.3	27.1	0.6
Divorce	*Estimate*	-0.0143	-0.0309	-0.0122
	Standard Error	(0.0022)	(0.0087)	(0.0032)
	Counterfactual Mean	0.11	0.17	0.09
	%Δ	-12.4	-18.1	-13.4

Notes: This table presents estimates of the mean effect per dollar of lper dollar of unearned income on the propensity to enter or leave marriage. The estimation sample is restricted to winners and not-yet winners that are tax filers. We use the delta method to calculate standard errors (reported in parenthesis), clustering on winner. To ease interpretability, each estimate is scaled by $10,000.

E Appendix: Analyses based on Employment Shocks from Firm Entry and Exit

E.1 Explanation of the Empirical Approach

Consider the following regression equation:

$$\log y_{i,t} - \log y_{i,t-1} = \beta_0 + \beta_1 X_{cz(i),t} + \gamma K_{i,t} + \epsilon_{i,t}, \tag{15}$$

where i is the worker or firm, y is an outcome such as the log earnings, $cz(i)$ is its commuting zone, $X_{cz,t}$ denotes the *growth* in the employment share by foreign-owned firms in that commuting zone, and $K_{i,t}$ is a vector of controls such as industry-year fixed effects, CZ fixed effects, and a polynomial in age. The parameter of interest is β_1, which captures how the entry of a foreign-owned firm affects a worker outcome or a firm outcome at a non-foreign-owned firm in the same commuting zone. We will refer to β_1 as the "indirect effect" of firm entry. The de-identified tax records linking workers to firms analyzed here allow for the construction of each of these variables, using the filing of Form 5472 as an indicator for foreign-ownership status. Basic descriptive statistics on the counts and outcome means of foreign-owned and non-foreign-owned firms for the 2015 cross-section are presented in Appendix Table A.25. When aggregating the number and share of employees at foreign-owned firms nationally over time, this data matches well the national trends reported by the BEA, as demonstrated in Appendix Figure A.24.

To identify β_1, we adapt the identification strategy common in the literature about the effects of immigration on non-immigrants in the same region (see Card, 2001). This literature uses the fact that immigrants cluster into regions in the US based on country of origin. To adapt this instrument to identify the effects of foreign-owned firm entry on workers, we first notice that employment at foreign-owned firms tends to be clustered by region and country of origin. For example, German-owned firms disproportionately employ workers in South Carolina in 2010 if they do so in 2005. This is analogous to the clustering of immigrants into regions. We construct a "Bartik instrument" as the predicted change in employment at, for example, German-owned firms in South Carolina between 2009 and 2010 using only information about (i) the share of workers at German-owned firms in South Carolina in 2005, and (ii) the change in aggregate employment by German-owned firms in *any other* region in the US between 2009 and 2010. Since this instrument is not formed using information about the change in employment by German-owned firms in South Carolina between 2009 and 2010, it does not depend directly on changes in South Carolina's business climate between 2009 and 2010. For example, it does not depend directly on infrastructure investments, improved educational opportunities, or changes in the generosity of tax incentives in South Carolina between 2009 and 2010. In Appendix Figure A.25, we provide evidence that ownership shares vary substantially over time and that spatial clustering occurs by nationality.[5]

[5]Spatial clustering of foreign investment by nationality in the US has also been recognized by Burchardi et al. (2016). They find that the stock of past migrants in a county from a certain origin country can help to predict today's foreign investment in the county by firms from that origin country.

Formally the identification challenge arises because, for example, the unobserved component of earnings growth $\epsilon_{i,t}$ (e.g., changes in the local demand for labor) may be correlated with $X_{cz,t}$. In order to form the instrument, we use the tax data on the firm's country of foreign ownership to construct the share $S^o_{cz,t}$ of all employment in commuting zone cz at firms whose owners are located in origin country o, defined by,

$$S^o_{cz,t} \equiv \frac{N^{F_o}_{cz,t}}{\sum_{cz'} N^{F_o}_{cz',t}} \tag{16}$$

Analogous to Card, 2001 and the subsequent immigration literature, we then construct the instrumental variable $Z_{cz,t}$ as,

$$Z_{cz,t} = \frac{\sum_o (\sum_{cz' \neq cz} N^{F_o}_{cz',t} - N^{F_o}_{cz',t-1}) S^o_{cz,t-5}}{N^F_{cz,t-5} + N^N_{cz,t-5}} \tag{17}$$

This is interpreted as the prediction of $X_{cz,t}$, formed only from the share of employment by firms from country o in cz dated at $t-5$ and the change in aggregate employment by o in the US from $t-1$ to t. Note that we modify the approach from the immigration literature slightly by leaving out own-commuting zone employment when constructing the aggregate change from $t-1$ to t, which helps to rule out confounding factors. The denominator is the total number of FTE workers in the commuting zone 5 years ago. Because $Z_{cz,t}$ is not a function of cz-specific changes between $t-1$ and t, it should satisfy that $Z_{cz,t}$ and the unexplained component of earnings growth are orthogonal (conditional on observed covariates that explain earnings growth). However, we see three major threats to identification, which we discuss below.

First, the instruments include the past share of employment at foreign-owned firms from various source countries as well as the change in the employment at such firms in other regions. This raises the concern that there may be regional shocks that are correlated with our instrument. For example, regions near the Canadian border may be affected also by trade shocks originating in Canada, that are correlated with the instrument. To deal with this concern, we include region-year fixed effects (where we define region as Census Division) in the regressions, that absorb all contemporaneous effects at the regional level.

Second, recent work by Jaeger et al. (2018) suggests that in the context of immigration past share of immigrants could have a direct effect on contemporaneous outcomes, if the adjustment to former immigrant waves is delayed. The analogous concern in our setting is that adjustment to past investment by foreign-owned firms is ongoing. Since our instrument leverages variation in the country of origin of foreign-owned firm investment across commuting zones, we can include as a control variable the share of past employment at foreign-owned firms (not separated by country of origin) in the commuting zone. We provide robustness results to our main regressions when also including the share of employment at foreign-owned firms in $t-5$ as a control variable. We find that our main results are quantitatively robust to adding this control.

The third threat to identification is that industry shocks may be correlated with the instrument. For example, German- or Japanese-owned firms may be more likely to be in the car industry and select commuting zones that are also abundant with other car industry firms. To deal with this concern, we also include fine industry-year fixed effects that absorb any contemporaneous nation-wide growth trends by industry. Furthermore, note that by including CZ-fixed effects in our first difference specification, we control for a CZ-specific linear time trend in the outcome variable.

In Appendix E.2 below, we present and discuss the parameter estimates and the results from a number of specifications and robustness checks. As shown in Appendix Table A.26, we find positive and statistically significant effects of a firm entry shock on existing firms for hiring of new employees, total wage bill growth within the firm, and value added growth. To put the estimates in context, we find that, if a firm employing 10% of employees in the commuting zone exits and lays off its workers, then workers in existing employment relationships at other firms in the same commuting zone experience a 5.3% decline in employment, a 6.3% decline in wage bill, a 9.6% decline in the firm's value added, and a 1.5% decline in wages of continuing workers. The effects are statistically significant when using standard errors clustered by commuting zone or clustered by country of origin. The effects are especially concentrated in large firms and in firms in the tradeables sector. As shown in Appendix Table A.27, we find larger effects on continuing workers among the top quintile group, whose average annual earnings are \$156,700, and the second highest quintile group, whose average annual earnings are \$63,200.

We consider several robustness and placebo checks. First, Appendix Table A.28 demonstrates that the results are broadly similar when considering the controls and alternate samples motivated by the potential identification threats discussed above. Second, Appendix Table A.29 demonstrates that the effects would not be statistically significant if measuring the outcomes in the pre-period as a placebo test. Third, Appendix Table A.30 demonstrates robustness to controlling for different covariates; this is true when using OLS or IV estimation. Fourth, Appendix Table A.31 demonstrates robustness to the leave-out definition, leaving out radii from 50 miles to 300 miles around the firm's own commuting zone.

Finally, in Appendix Table A.32, we consider effects of firm entry or exit on household income and tax outcomes for workers at other firms. To interpret the results, if a firm employing 10% of employees in the commuting zone exits and lays off its workers, then workers at existing firms in the same commuting zone experience a 4.5% decrease in household income, 4.0% decrease in net household income, and a 6.8% decrease in tax payments. Respectively, these results suggest that household responses provide little insurance against these shocks, the Federal tax-and-transfer system provides about a 9% rate of insurance against these shocks, and tax payments are more sensitive to these shocks than wages, which is consistent with a progressive tax system. Finally, we find that the layoff shock leads workers in other firms to claim EITC at a higher rate (extensive margin of EITC take-up) and to claim higher EITC deductions (intensive margin of EITC take-up).

In order to better understand the mechanisms through which these firm entry shocks are passed from foreign-owned firms to other firms in the same commuting zone, we investigate the role of workers moving between these two types of firms as a potential transmission mechanism. First, Appendix Table A.33 shows the differences in mean log earnings, mean firm effects, and mean worker effects provided to the workers at the foreign-owned firms relative to non-foreign-owned firms, using the two-way fixed effects estimates from the movers analyses presented in Section 4.2. It presents the results both overall and conditioning on firms of approximately the same size. The higher firm premiums (7.2%) and higher worker quality (12.7%) at these firms, as well as the noticeably higher firm premiums to higher-skilled workers demonstrated when using the estimator with firm-worker interactions in Figure A.26, are consistent with these firms having higher average productivity and stronger skill-augmenting technology. Since 87.1% of new hires at foreign firms from domestic firms are from the same commuting zone, these firm premiums are primarily paid to incumbents in the region. Furthermore, Appendix Table A.34 provides estimates for the largest origin countries in the sample, finding evidence that distance serves as a cost to entry so that average productivity is higher when the country of origin is further away. Second, as illustrated in Appendix Figure A.27 and supported by the regression coefficients in Appendix Table A.35, workers see an increase in wages when moving from non-foreign-owned to foreign-owned firms. This is true for raw and residual earnings measures, and it is true for the full sample as well as the sample of workers whose firms experienced a layoff with 30% or more of workers moving. While these findings are only suggestive, they are consistent with technological spillovers hypothesized in the literature on firm entry, in which workers (especially higher-skilled workers) learn new technological processes from new high-productivity firms and later communicate this technology to other firms in the same labor market.

E.2 Data Description and Results on Effects of Firm Entry and Exit

	Domestic	Foreign
Firms in Main Sample of Firms (thousands)	2,781.1	30.3
Firm-Location Pairs in Main Sample of Firms (thousands)	4,762.9	218.7
Number of Workers at Main Sample of Firms (millions):		
All Workers:	77.1	5.2
FTE Analysis Sample:	41.3	3.6
Mean Wage at Main Sample of Firms (thousands):		
All Workers:	41.4	60.7
FTE Analysis Sample:	62.6	75.7
Value Added per Worker at Main Sample of Firms (thousands):		
All Workers:	82.7	153.1
FTE Analysis Sample:	154.3	220.1
Sample Exit Rates among those Active in 2014:		
Workers:	0.271	0.217
Firms:	0.123	0.137

Table A.25: 2015 Cross-sectional Descriptive Statistics on the Samples used in the Firm Entry Analysis

Notes: This table displays descriptive statistics for domestic-owned and foreign-owned firms that file forms 1120, 1120-S, and 1065, matched to subsidiaries and W-2 forms.

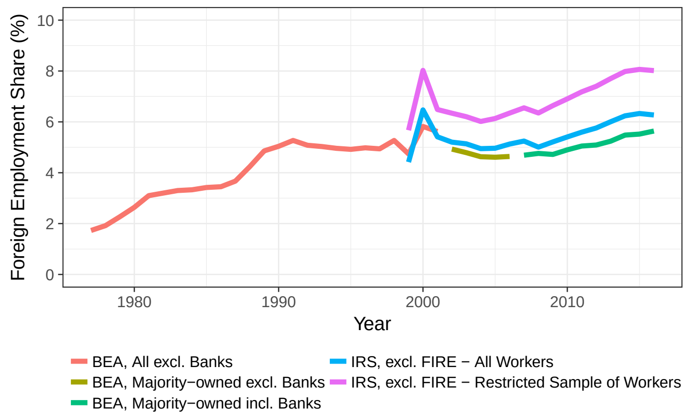
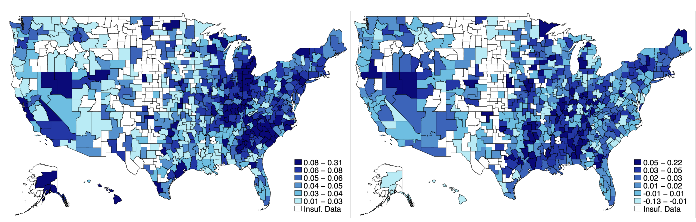
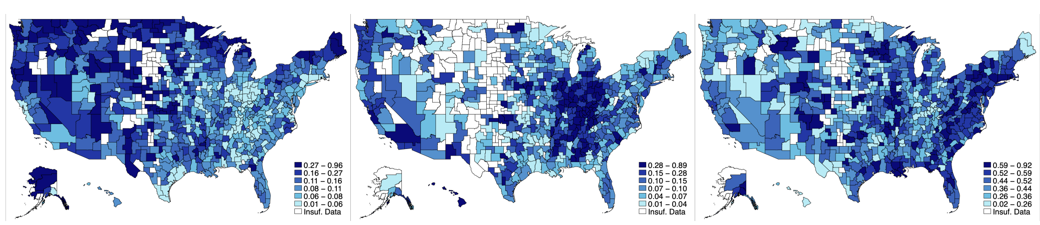


Figure A.24: Validating that Foreign Ownership is Measured Similarly to BEA

Notes: This figure compares the total employment counts and shares at foreign-owned firms over time observed in tax records to those reported by the BEA, where BEA has provided three distinct series of measurements over time.



(a) Any Owner Country, 2001 (b) Any Owner Country, 2001-2015 Change



(c) Canadian Owned (d) East Asian Owned (e) Western European Owned

Figure A.25: Geographic Concentration and Changes in Employment Shares at Foreign-owned Firms

Notes: This figure compares shares of employment at foreign-owned firms for all countries of ownership in 2001 and the change from 2001 to 2015, as well as disaggregating by groups of countries of ownership and averaging across all years.

	Full Sample	By Firm Size			By Sector	
		Size 1-9	Size 10-99	Size 100+	Tradables	Non-tradables
Panel A.	Outcome: Log Change in Value Added					
Second-Stage Coefficient	**0.96**	**0.12**	**0.54**	**2.66**	**3.38**	**0.50**
(Std. Error Clustered by Commuting Zone)	(0.30)	(0.10)	(0.18)	(1.14)	(1.56)	(0.23)
(Std. Error Clustered by Country of Origin)	(0.51)	(0.08)	(0.20)	(1.64)	(3.10)	(0.23)
First-Stage Coefficient	0.56	0.59	0.53	0.49	0.53	0.48
(F-statistic Clustered by Commuting Zone)	(232)	(361)	(241)	(112)	(128)	(143)
(F-statistic Clustered by Country of Origin)	(42)	(40)	(52)	(65)	(46)	(51)
Number of Firms by Commuting Zones (Millions)	41.8	34.9	6.5	0.5	6.0	6.0
Number of Workers (Millions, measured at $t-1$)	416.8	96.2	158.5	162.2	98.3	63.3
Panel B.	Outcome: Log Change in Employment					
Second-Stage Coefficient	**0.53**	**0.02**	**0.40**	**1.55**	**1.22**	**0.72**
(Std. Error Clustered by Commuting Zone)	(0.14)	(0.08)	(0.16)	(0.52)	(0.43)	(0.25)
(Std. Error Clustered by Country of Origin)	(0.18)	(0.07)	(0.17)	(0.54)	(0.43)	(0.26)
First-Stage Coefficient	0.56	0.59	0.53	0.50	0.53	0.48
(F-statistic Clustered by Commuting Zone)	(235)	(364)	(246)	(119)	(130)	(143)
(F-statistic Clustered by Country of Origin)	(44)	(39)	(52)	(66)	(49)	(53)
Number of Firms by Commuting Zones (Millions)	46.0	38.3	7.1	0.5	6.4	6.4
Number of Workers (Millions, measured at $t-1$)	477.3	105.1	175.8	196.5	107.3	71.1
Panel C.	Outcome: Log Change in Wage Bill					
Second-Stage Coefficient	**0.63**	**0.00**	**0.41**	**1.62**	**1.42**	**1.19**
(Std. Error Clustered by Commuting Zone)	(0.17)	(0.10)	(0.18)	(0.53)	(0.47)	(0.35)
(Std. Error Clustered by Country of Origin)	(0.22)	(0.10)	(0.19)	(0.56)	(0.50)	(0.36)
First-Stage Coefficient	0.56	0.59	0.53	0.50	0.53	0.48
(F-statistic Clustered by Commuting Zone)	(235)	(364)	(246)	(119)	(130)	(143)
(F-statistic Clustered by Country of Origin)	(44)	(39)	(52)	(66)	(49)	(53)
Number of Firms by Commuting Zones (Millions)	46.0	38.3	7.1	0.5	6.4	6.4
Number of Workers (Millions, measured at $t-1$)	477.3	105.1	175.8	196.5	107.3	71.1
Panel D.	Outcome: Log Change in Earnings of Continuing Workers					
Second-Stage Coefficient	**0.15**	**0.01**	**0.06**	**0.40**	**0.39**	**0.19**
(Std. Error Clustered by Commuting Zone)	(0.07)	(0.05)	(0.07)	(0.14)	(0.16)	(0.09)
(Std. Error Clustered by Country of Origin)	(0.08)	(0.07)	(0.08)	(0.15)	(0.17)	(0.09)
First-Stage Coefficient	0.56	0.59	0.54	0.50	0.53	0.49
(F-statistic Clustered by Commuting Zone)	(239)	(367)	(249)	(123)	(134)	(149)
(F-statistic Clustered by Country of Origin)	(44)	(39)	(52)	(66)	(48)	(53)
Number of Firms by Commuting Zones (Millions)	44.6	37.0	7.1	0.5	6.3	6.2
Number of Workers (Millions, measured at $t-1$)	369.6	83.4	130.9	155.3	87.2	54.4

Table A.26: Firm-level Bartik IV Estimates: Main Effects on All Other Firms

Notes: The outcome sample only includes continuing domestic firms. Observations are weighted by lagged firm size. Controls are industry-year indicators, Census-division-year indicators, measures of urban concentration, and the sum of commuting zone exposure shares.

		By Earnings Quintile Group				
	Full Sample	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Outcome: Log Change in Earnings of Continuing Workers					
Second-Stage Coefficient	**0.15**	**0.06**	**0.04**	**0.08**	**0.27**	**0.32**
(Std. Error Clustered by Commuting Zone)	(0.07)	(0.12)	(0.09)	(0.08)	(0.09)	(0.12)
(Std. Error Clustered by Country of Origin)	(0.08)	(0.13)	(0.12)	(0.09)	(0.10)	(0.12)
First-Stage Coefficient	0.56	0.55	0.56	0.56	0.56	0.55
(F-statistic Clustered by Commuting Zone)	(239)	(235)	(237)	(238)	(238)	(238)
(F-statistic Clustered by Country of Origin)	(44)	(50)	(47)	(47)	(46)	(47)
Number of Firms by Commuting Zones (Millions)	44.6	20.1	19.6	18.8	17.0	16.1
Number of Workers (Millions, measured at $t-1$)	369.6	73.9	73.9	73.9	73.9	73.9

Table A.27: Worker-level Bartik IV Estimates: Indirect Effects of Firm Entry on Continuing Workers at Non-foreign-owned Firms in the Same Commuting Zone

Notes: Controls are 3-digit-industry-year fixed effects, Census Division-year fixed effects, CZ fixed effects, and a polynomial in worker age. Recall that firm fixed effects are removed through the first-differenced specification. Standard errors are clustered at the CZ-year level. The sample only includes continuing workers at domestic-owned firms. We divide workers into five wage quintile groups within each CZ-year based on the ordering of their lagged wages.

	(1) Baseline	(2) Control Lag IV	(3) Control NAICS-6	(4) Leave out 300m Radius	(5) Exclude Dom MNE	(6) Exclude Tax Havens	(7) DHS Transform
Panel A.	**Outcome: Log Change in Value Added**						
Second-Stage Coefficient	**0.96**	**0.98**	**0.93**	**0.97**	**0.87**	**1.10**	**1.44**
(Std. Error Clustered by Commuting Zone)	(0.30)	(0.32)	(0.26)	(0.35)	(0.25)	(0.36)	(0.47)
(Std. Error Clustered by Country of Origin)	(0.51)	(0.51)	(0.34)	(0.51)	(0.35)	(0.62)	(0.53)
First-Stage Coefficient	0.56	0.54	0.56	0.59	0.57	0.55	0.57
(F-statistic Clustered by Commuting Zone)	(232)	(208)	(233)	(143)	(260)	(233)	(264)
(F-statistic Clustered by Country of Origin)	(42)	(42)	(43)	(42)	(40)	(35)	(41)
Number of Firms by Commuting Zones (Millions)	41.8	40.5	41.8	41.8	40.6	41.8	66.6
Number of Workers (Millions, measured at $t-1$)	416.8	401.0	416.8	416.8	344.1	416.8	497.8
Effective Number of Country Shocks (Inverse HHI)	153.6	153.6	153.6	153.6	153.6	122.2	153.6
Panel B.	**Outcome: Log Change in Employment**						
Second-Stage Coefficient	**0.53**	**0.56**	**0.50**	**0.56**	**0.42**	**0.55**	**0.83**
(Std. Error Clustered by Commuting Zone)	(0.14)	(0.15)	(0.14)	(0.17)	(0.14)	(0.17)	(0.29)
(Std. Error Clustered by Country of Origin)	(0.18)	(0.18)	(0.17)	(0.18)	(0.18)	(0.22)	(0.23)
First-Stage Coefficient	0.56	0.54	0.56	0.59	0.57	0.55	0.58
(F-statistic Clustered by Commuting Zone)	(235)	(211)	(236)	(145)	(258)	(241)	(270)
(F-statistic Clustered by Country of Origin)	(44)	(43)	(44)	(44)	(42)	(37)	(41)
Number of Firms by Commuting Zones (Millions)	46.0	44.6	46.0	46.0	44.6	46.0	69.2
Number of Workers (Millions, measured at $t-1$)	477.3	459.6	477.3	477.3	395.6	477.3	519.7
Effective Number of Country Shocks (Inverse HHI)	153.6	153.6	153.6	153.6	153.6	122.2	153.6
Panel C.	**Outcome: Log Change in Wage Bill**						
Second-Stage Coefficient	**0.63**	**0.61**	**0.59**	**0.69**	**0.49**	**0.67**	**0.88**
(Std. Error Clustered by Commuting Zone)	(0.17)	(0.19)	(0.17)	(0.21)	(0.17)	(0.20)	(0.29)
(Std. Error Clustered by Country of Origin)	(0.22)	(0.22)	(0.21)	(0.22)	(0.22)	(0.27)	(0.26)
First-Stage Coefficient	0.56	0.54	0.56	0.59	0.57	0.55	0.58
(F-statistic Clustered by Commuting Zone)	(235)	(211)	(236)	(145)	(258)	(241)	(270)
(F-statistic Clustered by Country of Origin)	(44)	(43)	(44)	(44)	(42)	(37)	(41)
Number of Firms by Commuting Zones (Millions)	46.0	44.6	46.0	46.0	44.6	46.0	69.2
Number of Workers (Millions, measured at $t-1$)	477.3	459.6	477.3	477.3	395.6	477.3	519.7
Effective Number of Country Shocks (Inverse HHI)	153.6	153.6	153.6	153.6	153.6	122.2	153.6
Panel D.	**Outcome: Log Change in Earnings of Continuing Workers**						
Second-Stage Coefficient	**0.15**	**0.10**	**0.13**	**0.17**	**0.09**	**0.15**	**0.15**
(Std. Error Clustered by Commuting Zone)	(0.07)	(0.08)	(0.07)	(0.09)	(0.07)	(0.08)	(0.07)
(Std. Error Clustered by Country of Origin)	(0.08)	(0.08)	(0.07)	(0.08)	(0.07)	(0.09)	(0.08)
First-Stage Coefficient	0.56	0.55	0.56	0.59	0.57	0.56	0.56
(F-statistic Clustered by Commuting Zone)	(239)	(214)	(240)	(150)	(265)	(247)	(239)
(F-statistic Clustered by Country of Origin)	(44)	(43)	(44)	(44)	(41)	(37)	(44)
Number of Firms by Commuting Zones (Millions)	44.6	43.3	44.6	44.6	43.3	44.6	44.6
Number of Workers (Millions, measured at $t-1$)	369.6	356.0	369.6	369.6	304.3	369.6	369.6
Effective Number of Country Shocks (Inverse HHI)	153.6	153.6	153.6	153.6	153.6	122.2	153.6

Table A.28: Firm-level Bartik IV Estimates: Robustness Checks

Notes: The outcome sample only includes continuing domestic firms (unless otherwise specified). Observa tions are weighted by lagged firm size (unless otherwise specified). Controls are industry-year indicators, Census-division-year indicators, measures of urban concentration, and the sum of commuting zone exposure shares (unless otherwise specified).

	Value Added		Employment		Wage bill		Earnings of Cont. Workers	
	Main	Placebo	Main	Placebo	Main	Placebo	Main	Placebo
Second-Stage:								
Coefficient	**0.96**	**-0.05**	**0.53**	**-0.17**	**0.63**	**-0.09**	**0.15**	**0.04**
(Std. Error Clustered by Commuting Zone)	(0.30)	(0.22)	(0.14)	(0.12)	(0.17)	(0.16)	(0.07)	(0.09)
(Std. Error Clustered by Country of Origin)	(0.51)	(0.45)	(0.18)	(0.14)	(0.22)	(0.16)	(0.08)	(0.09)
First-Stage:								
Coefficient	0.56	0.66	0.56	0.66	0.56	0.66	0.56	0.67
(F-statistic Clustered by Commuting Zone)	(232)	(341)	(235)	(351)	(235)	(351)	(239)	(360)
(F-statistic Clustered by Country of Origin)	(42)	(27)	(44)	(27)	(44)	(27)	(44)	(26)
Number of Firms by Commuting Zones (Millions)	41.8	36.2	46.0	38.7	46.0	38.7	44.6	37.6
Number of Workers (Millions, measured at $t-1$)	416.8	402.0	477.3	441.1	477.3	441.1	369.6	336.8

Table A.29: Firm-level Bartik IV Estimates: Placebo Test

Notes: The outcome sample only includes continuing domestic firms. Observations are weighted by lagged firm size. Controls are industry-year indicators, Census-division-year indicators, measures of urban concentration, and the sum of commuting zone exposure shares. Placebo outcomes are measured as changes between t_0-2 and t_0-1, where t_0 is the time period at which the exposure shares are measured.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Control Specification									
CZ-year domestic employment share	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry-year fixed effects	✗	✓	✓	✓	✓	✓	✓	✓	✓
Census-division-year fixed effects	✗	✗	✓	✓	✓	✓	✓	✓	✓
CZ controls:									
Urban density measures (pre-period)	✗	✗	✗	✓	✓	✓	✓	✓	✓
Educational attainment measures (pre-period)	✗	✗	✗	✗	✓	✓	✓	✓	✓
Poverty and employment measures (pre-period)	✗	✗	✗	✗	✗	✓	✓	✓	✓
Farm and manufacturing measures (pre-period)	✗	✗	✗	✗	✗	✗	✓	✓	✓
CZ-year domestic employment share × Financial Crisis	✗	✗	✗	✗	✗	✗	✗	✓	✓
CZ-year domestic employment share × All 3-year intervals	✗	✗	✗	✗	✗	✗	✗	✗	✓
Panel A.	**Outcome: Log Change in Value Added**								
OLS Coefficient	**0.05**	**0.49**	**0.34**	**0.33**	**0.33**	**0.32**	**0.32**	**0.31**	**0.31**
(Std. Error Clustered by Commuting Zone)	(0.13)	(0.10)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Second-Stage Coefficient	**-0.26**	**0.66**	**1.01**	**0.96**	**0.93**	**0.90**	**0.90**	**0.88**	**1.13**
(Std. Error Clustered by Commuting Zone)	(0.15)	(0.27)	(0.32)	(0.30)	(0.30)	(0.30)	(0.30)	(0.31)	(0.49)
(Std. Error Clustered by Country of Origin)	(0.37)	(0.44)	(0.50)	(0.51)	(0.50)	(0.49)	(0.49)	(0.48)	(0.90)
First-Stage Coefficient	0.96	0.61	0.56	0.56	0.56	0.56	0.56	0.55	0.44
(F-statistic Clustered by Commuting Zone)	(2,890)	(232)	(232)	(232)	(233)	(232)	(231)	(224)	(100)
(F-statistic Clustered by Country of Origin)	(1,367)	(30)	(42)	(42)	(43)	(42)	(42)	(40)	(19)
Number of Firms by Commuting Zones (Millions)	41.8	41.8	41.8	41.8	41.8	41.8	41.8	41.8	41.8
Number of Workers (Millions, measured at $t-1$)	416.8	416.8	416.8	416.8	416.8	416.8	416.6	416.6	416.6
Panel B.	**Outcome: Log Change in Employment**								
OLS Coefficient	**-0.03**	**0.28**	**0.24**	**0.22**	**0.22**	**0.21**	**0.21**	**0.21**	**0.22**
(Std. Error Clustered by Commuting Zone)	(0.07)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Second-Stage Coefficient	**-0.24**	**0.46**	**0.60**	**0.53**	**0.50**	**0.46**	**0.46**	**0.46**	**0.66**
(Std. Error Clustered by Commuting Zone)	(0.07)	(0.15)	(0.15)	(0.14)	(0.13)	(0.13)	(0.13)	(0.13)	(0.21)
(Std. Error Clustered by Country of Origin)	(0.18)	(0.20)	(0.22)	(0.18)	(0.18)	(0.17)	(0.17)	(0.17)	(0.30)
First-Stage Coefficient	0.95	0.61	0.56	0.56	0.56	0.56	0.56	0.56	0.44
(F-statistic Clustered by Commuting Zone)	(2,890)	(233)	(235)	(235)	(236)	(235)	(234)	(227)	(102)
(F-statistic Clustered by Country of Origin)	(1,402)	(30)	(44)	(44)	(44)	(44)	(43)	(42)	(19)
Number of Firms by Commuting Zones (Millions)	46.0	46.0	46.0	46.0	46.0	46.0	45.9	45.9	45.9
Number of Workers (Millions, measured at $t-1$)	477.3	477.3	477.3	477.3	477.3	477.3	477.1	477.1	477.1
Panel C.	**Outcome: Log Change in Wage Bill**								
OLS Coefficient	**-0.58**	**0.39**	**0.31**	**0.29**	**0.30**	**0.29**	**0.29**	**0.28**	**0.29**
(Std. Error Clustered by Commuting Zone)	(0.13)	(0.07)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Second-Stage Coefficient	**-1.19**	**0.54**	**0.70**	**0.63**	**0.59**	**0.55**	**0.54**	**0.53**	**0.72**
(Std. Error Clustered by Commuting Zone)	(0.14)	(0.18)	(0.18)	(0.17)	(0.16)	(0.16)	(0.16)	(0.16)	(0.26)
(Std. Error Clustered by Country of Origin)	(0.36)	(0.25)	(0.25)	(0.22)	(0.22)	(0.21)	(0.20)	(0.20)	(0.35)
First-Stage Coefficient	0.95	0.61	0.56	0.56	0.56	0.56	0.56	0.56	0.44
(F-statistic Clustered by Commuting Zone)	(2,890)	(233)	(235)	(235)	(236)	(235)	(234)	(227)	(102)
(F-statistic Clustered by Country of Origin)	(1,402)	(30)	(44)	(44)	(44)	(44)	(43)	(42)	(19)
Number of Firms by Commuting Zones (Millions)	46.0	46.0	46.0	46.0	46.0	46.0	45.9	45.9	45.9
Number of Workers (Millions, measured at $t-1$)	477.3	477.3	477.3	477.3	477.3	477.3	477.1	477.1	477.1
Panel D.	**Outcome: Log Change in Earnings of Cont. Workers**								
OLS Coefficient	**-0.68**	**0.16**	**0.13**	**0.13**	**0.13**	**0.13**	**0.13**	**0.12**	**0.13**
(Std. Error Clustered by Commuting Zone)	(0.09)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Second-Stage Coefficient	**-1.21**	**0.11**	**0.17**	**0.15**	**0.13**	**0.13**	**0.13**	**0.11**	**0.13**
(Std. Error Clustered by Commuting Zone)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.11)
(Std. Error Clustered by Country of Origin)	(0.23)	(0.11)	(0.09)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.13)
First-Stage Coefficient	0.95	0.61	0.56	0.56	0.56	0.56	0.56	0.56	0.44
(F-statistic Clustered by Commuting Zone)	(2,930)	(238)	(239)	(239)	(240)	(239)	(238)	(231)	(104)
(F-statistic Clustered by Country of Origin)	(1,340)	(29)	(43)	(44)	(44)	(43)	(43)	(41)	(19)
Number of Firms by Commuting Zones (Millions)	44.6	44.6	44.6	44.6	44.6	44.6	44.6	44.6	44.6
Number of Workers (Millions, measured at $t-1$)	369.6	369.6	369.6	369.6	369.6	369.6	369.5	369.5	369.5

Table A.30: Firm-level Bartik IV Estimates: Robustness to Alternative Control Sets

Notes: The outcome sample only includes continuing domestic firms. Observations are weighted by lagged firm size. Controls are indicated at the top of the table. Our baseline control set is in column (4).

	Leave out No CZ (include Own)	Leave out Own CZ	Leave out CZs within Radius (based on nearest distance in miles)						Leave out Entire Census Division
			50	100	150	200	250	300	
	Number of CZs excluded								
Mean	0	1	7	16	28	42	59	76	77
25th quantile	0	1	5	10	17	24	30	37	58
50th quantile	0	1	8	16	28	42	57	74	84
75th quantile	0	1	9	21	38	59	84	114	104
Panel A.	**Outcome: Log Change in Value Added**								
Second-Stage Coefficient	**0.87**	**0.96**	**0.96**	**0.94**	**0.95**	**0.97**	**1.01**	**0.97**	**0.89**
(Std. Error Clustered by Commuting Zone)	(0.28)	(0.30)	(0.32)	(0.33)	(0.32)	(0.33)	(0.33)	(0.35)	(0.32)
(Std. Error Clustered by Country of Origin)	(0.51)	(0.51)	(0.51)	(0.51)	(0.51)	(0.51)	(0.51)	(0.51)	(0.51)
First-Stage Coefficient	0.59	0.56	0.56	0.56	0.58	0.58	0.59	0.59	0.61
(F-statistic Clustered by Commuting Zone)	(273)	(232)	(212)	(199)	(185)	(172)	(161)	(143)	(189)
(F-statistic Clustered by Country of Origin)	(42)	(42)	(42)	(42)	(42)	(42)	(42)	(42)	(42)
Number of Firms by Commuting Zones (Millions)	41.8	41.8	41.8	41.8	41.8	41.8	41.8	41.8	41.8
Number of Workers (Millions, measured at $t-1$)	416.8	416.8	416.8	416.8	416.8	416.8	416.8	416.8	416.8
Panel B.	**Outcome: Log Change in Employment**								
Second-Stage Coefficient	**0.52**	**0.53**	**0.54**	**0.55**	**0.55**	**0.55**	**0.57**	**0.56**	**0.55**
(Std. Error Clustered by Commuting Zone)	(0.13)	(0.14)	(0.15)	(0.15)	(0.15)	(0.16)	(0.16)	(0.17)	(0.15)
(Std. Error Clustered by Country of Origin)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
First-Stage Coefficient	0.59	0.56	0.56	0.56	0.58	0.58	0.59	0.59	0.61
(F-statistic Clustered by Commuting Zone)	(277)	(235)	(214)	(201)	(188)	(175)	(163)	(145)	(192)
(F-statistic Clustered by Country of Origin)	(44)	(44)	(44)	(44)	(44)	(44)	(44)	(44)	(44)
Number of Firms by Commuting Zones (Millions)	46.0	46.0	46.0	46.0	46.0	46.0	46.0	46.0	46.0
Number of Workers (Millions, measured at $t-1$)	477.3	477.3	477.3	477.3	477.3	477.3	477.3	477.3	477.3
Panel C.	**Outcome: Log Change in Wage Bill**								
Second-Stage Coefficient	**0.61**	**0.63**	**0.64**	**0.64**	**0.65**	**0.66**	**0.69**	**0.69**	**0.64**
(Std. Error Clustered by Commuting Zone)	(0.16)	(0.17)	(0.18)	(0.18)	(0.19)	(0.19)	(0.20)	(0.21)	(0.18)
(Std. Error Clustered by Country of Origin)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)
First-Stage Coefficient	0.59	0.56	0.56	0.56	0.58	0.58	0.59	0.59	0.61
(F-statistic Clustered by Commuting Zone)	(277)	(235)	(214)	(201)	(188)	(175)	(163)	(145)	(192)
(F-statistic Clustered by Country of Origin)	(44)	(44)	(44)	(44)	(44)	(44)	(44)	(44)	(44)
Number of Firms by Commuting Zones (Millions)	46.0	46.0	46.0	46.0	46.0	46.0	46.0	46.0	46.0
Number of Workers (Millions, measured at $t-1$)	477.3	477.3	477.3	477.3	477.3	477.3	477.3	477.3	477.3
Panel D.	**Outcome: Log Change in Earnings of Continuing Workers**								
Second-Stage Coefficient	**0.15**	**0.15**	**0.14**	**0.14**	**0.15**	**0.16**	**0.17**	**0.17**	**0.14**
(Std. Error Clustered by Commuting Zone)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)
(Std. Error Clustered by Country of Origin)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
First-Stage Coefficient	0.59	0.56	0.56	0.56	0.58	0.59	0.59	0.59	0.61
(F-statistic Clustered by Commuting Zone)	(281)	(239)	(218)	(205)	(192)	(179)	(167)	(150)	(196)
(F-statistic Clustered by Country of Origin)	(44)	(44)	(44)	(44)	(44)	(44)	(44)	(44)	(44)
Number of Firms by Commuting Zones (Millions)	44.6	44.6	44.6	44.6	44.6	44.6	44.6	44.6	44.6
Number of Workers (Millions, measured at $t-1$)	369.6	369.6	369.6	369.6	369.6	369.6	369.6	369.6	369.6

Table A.31: Firm-level Bartik IV Estimates: Robustness to Leave-out Distances

Notes: The outcome sample only includes continuing domestic firms. Observations are weighted by lagged firm size. Controls are industry-year indicators, Census-division-year indicators, measures of urban concentration, and the sum of commuting zone exposure shares. Our baseline specification is in column (2).

Outcome: (unit transformation)	**Earnings** (log)	**Income before Taxes** (log)	**Income after Taxes** (log)	**Tax Payments** (log)	**EITC Claims** (DHS transform)	**EITC Amount** (DHS transform)
2SLS Indirect Effect	0.466***	0.588***	0.538***	0.754***	-0.241	-0.312
	(0.138)	(0.149)	(0.144)	(0.232)	(0.182)	(0.203)
First Stage Coefficient	0.598***	0.597***	0.597***	0.600***	0.621***	0.621***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.030)	(0.030)
First Stage F-statistics	297	291	291	295	421	421
Firm Observations (Millions)	45.9	34.8	34.7	35.4	16.3	16.3

Table A.32: Firm-level Bartik IV Estimates: Various Income and Tax Measures

Notes: Controls are as in the main specifications for worker-weighted regressions described in the text. Standard errors are clustered at the CZ-year level. Firm-level observations are weighted by lagged firm size. The Davis et al. (1996, DHS) transform of x_t is given by $\frac{x_t - x_{t-1}}{(x_t + x_{t-1})/2}$ and is meant to approximate the log-difference when x has a large share of non-positive numbers and is thus not defined in the log transform. EITC claims and amounts are per-capita within the firm, and firms are equally weighted when using these per-capita outcomes.

Size Bin:	All	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Foreign 2010-15:																
Log Earnings	0.195	0.356	0.257	0.294	0.243	0.223	0.194	0.155	0.134	0.105	0.078	0.089	0.083	0.100	0.075	0.100
Firm Effect	0.072	0.125	0.090	0.105	0.087	0.080	0.067	0.053	0.049	0.042	0.036	0.040	0.038	0.045	0.039	0.045
Worker Effect	0.128	0.235	0.177	0.200	0.164	0.151	0.134	0.109	0.091	0.068	0.047	0.051	0.047	0.055	0.039	0.058
Mulinational 2010-15:																
Log Earnings	0.236	0.409	0.324	0.259	0.238	0.215	0.165	0.168	0.144	0.132	0.123	0.104	0.120	0.136	0.183	0.209
Firm Effect	0.084	0.158	0.104	0.101	0.076	0.075	0.052	0.058	0.051	0.045	0.042	0.037	0.041	0.047	0.059	0.063
Worker Effect	0.155	0.236	0.228	0.158	0.169	0.143	0.114	0.116	0.097	0.093	0.087	0.072	0.085	0.095	0.130	0.152
Foreign 2001-06:																
Log Earnings	0.170	0.263	0.201	0.236	0.216	0.194	0.172	0.143	0.111	0.101	0.084	0.071	0.079	0.065	0.059	0.092
Firm Effect	0.067	0.096	0.073	0.084	0.079	0.072	0.066	0.056	0.046	0.046	0.041	0.040	0.041	0.034	0.035	0.048
Worker Effect	0.105	0.171	0.130	0.154	0.140	0.125	0.110	0.091	0.069	0.058	0.046	0.034	0.039	0.034	0.027	0.045
Foreign by Clusters:																
Firm Effect - 20	0.071	0.132	0.093	0.108	0.088	0.080	0.067	0.052	0.047	0.039	0.035	0.038	0.039	0.046	0.038	0.042
Firm Effect - 30	0.070	0.127	0.090	0.105	0.087	0.079	0.064	0.051	0.044	0.037	0.031	0.037	0.036	0.046	0.036	0.043
Firm Effect - 40	0.070	0.130	0.090	0.106	0.087	0.078	0.064	0.051	0.045	0.038	0.032	0.038	0.038	0.045	0.035	0.042
Firm Effect - 50	0.069	0.130	0.089	0.105	0.086	0.078	0.062	0.048	0.043	0.037	0.031	0.036	0.036	0.042	0.033	0.041
Worker Effect - 20	0.129	0.228	0.176	0.199	0.164	0.151	0.134	0.110	0.094	0.071	0.048	0.053	0.047	0.055	0.039	0.061
Worker Effect - 30	0.131	0.232	0.178	0.202	0.165	0.153	0.137	0.112	0.096	0.073	0.052	0.054	0.049	0.055	0.041	0.060
Worker Effect - 40	0.130	0.230	0.178	0.201	0.165	0.153	0.137	0.112	0.095	0.072	0.050	0.053	0.048	0.056	0.042	0.061
Worker Effect - 50	0.132	0.230	0.179	0.202	0.166	0.154	0.139	0.114	0.097	0.074	0.052	0.055	0.049	0.059	0.044	0.062

Table A.33: Mean Differences between Foreign-owned and Domestic-owned Firms by Size Bin

Notes: This table presents the mean difference in log value added, log wage bill per worker, log earnings, firm effects, and worker effects. Firms are grouped into log size bins, and means are computed within bins.

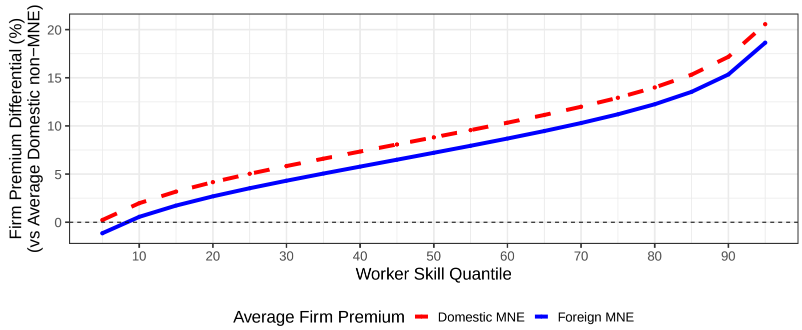


Figure A.26: Mean Differences between Foreign-owned and Domestic-owned Firms with Interacted Fixed Effects

Notes: This figure presents estimates of the model with interacted firm-worker fixed effects from the grouped fixed effect estimator during 2010-2015. The horizontal axis is a quantile in the distribution of estimated worker skill level. The vertical axis is the difference in the average firm premium for a worker of a given skill level for foreign (blue) or domestic (red) multinationals, relative to the average domestic non-multinational.

Ownership		Mean Difference in:			
	N	Log VA	Log Earnings	Firm Effects	Worker Effects
AE	0.009	0.442	0.155	0.044	0.117
AS	0.318	0.834	0.283	0.101	0.198
AU	0.145	1.395	0.269	0.108	0.160
BE	0.190	1.404	0.286	0.104	0.187
BR	0.204	0.416	0.172	0.055	0.112
CA	1.926	0.931	0.151	0.052	0.104
CH	0.088	-0.516	-0.065	-0.041	-0.017
CO	0.010	-0.300	0.052	0.030	0.025
DN	0.218	1.285	0.302	0.109	0.200
FI	0.125	1.382	0.306	0.117	0.193
FR	1.545	1.178	0.215	0.082	0.142
GM	2.312	1.255	0.226	0.090	0.139
HK	0.053	0.359	0.121	0.041	0.082
IN	0.176	0.733	0.139	0.051	0.106
IR	0.378	1.361	0.355	0.124	0.239
IS	0.130	0.808	0.257	0.097	0.179
IT	0.309	1.035	0.231	0.090	0.144
JA	2.795	1.470	0.175	0.079	0.098
LU	0.352	2.334	0.298	0.104	0.193
MX	0.190	-0.079	0.018	0.005	0.016
NL	1.196	1.700	0.280	0.101	0.182
NO	0.076	1.369	0.395	0.151	0.255
NZ	0.080	0.792	0.342	0.127	0.225
RS	0.020	-0.116	0.107	0.025	0.096
SF	0.017	0.351	0.157	0.057	0.103
SK	0.220	0.160	0.103	0.042	0.066
SN	0.106	0.789	0.175	0.054	0.132
SP	0.207	0.849	0.216	0.080	0.145
SW	0.664	1.301	0.303	0.115	0.191
SZ	1.576	1.671	0.277	0.105	0.174
TU	0.012	-0.152	0.147	0.059	0.090
TW	0.094	0.327	0.001	0.005	-0.002
UK	2.632	1.186	0.276	0.100	0.185
VZ	0.022	-0.365	0.041	0.012	0.028

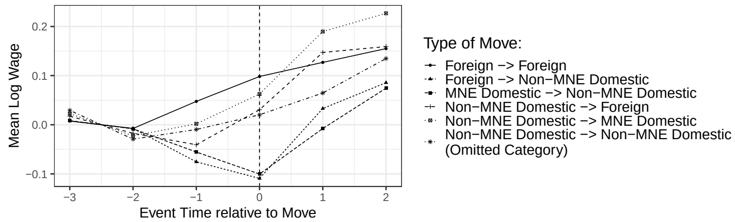
Table A.34: Differences between Foreign-owned and Domestic-owned Firms by Origin Country

Notes: This table presents the mean differences in log value added, log earnings, firm effects, and worker effects between the largest foreign ownership countries and the non-foreign-owned firms.

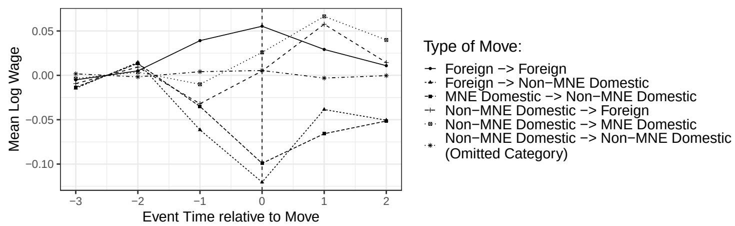
	(1)	(2)	(3)
Type of Move:			
Domestic to Foreign	0.078***	0.059***	-0.002
	(0.002)	(0.003)	(0.004)
	(N=242,207)	(N=126,178)	(N=48,795)
Foreign to Domestic	-0.056***	-0.052***	0.014***
	(0.002)	(0.004)	(0.004)
	(N=172,896)	(N=46,729)	(N=37,966)
Foreign to Foreign	0.020***	0.042***	0.006*
	(0.003)	(0.004)	(0.003)
	(N=246,192)	(N=128,396)	(N=246,192)
Domestic to Domestic	0	0	0
(Omitted Category)	(N=7,900,458)	(N=3,290,933)	(N=223,424)
Specification Details:			
Domestic Firms Restriction	Exclude MNE	Exclude MNE	Only include MNE
Type of Separation	All	Mass Layoff	All

Table A.35: Earnings growth for movers and stayers

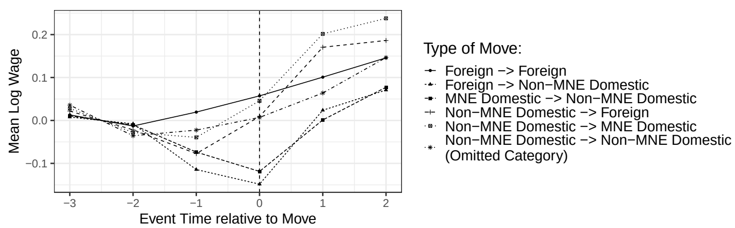
Notes: N denotes sample size. Standard errors in parentheses. The sample consists of only workers who were employed for two straight years at one firm followed by two straight years at a different firm.



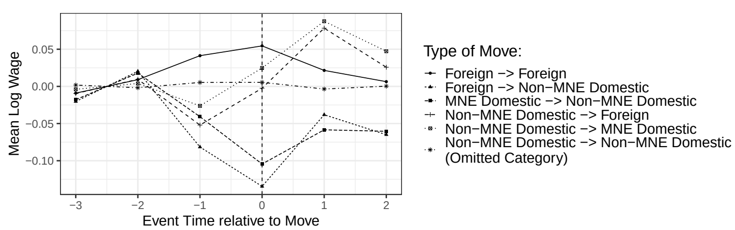
(a) Full sample, raw outcome



(b) Full sample, residual outcome



(c) Layoff sample, raw outcome



(d) Layoff sample, residual outcome

Figure A.27: Movers Event Study on Mean Earnings

Notes: In this figure, we compute the mean of the log earnings for various types of moves at event time zero, either when moving from a domestic-owned to a foreign-owned firm, or when moving from a foreign-owned to a domestic-owned firm. We compare these to workers who remain at the same domestic firm, as this is the omitted category in the regression for movers. We measure the outcome both as raw mean of log earnings, and residual mean of log earnings. We consider both the full sample of movers, and those who moved as part of a large layoff in which at least thirty percent of workers moved.

F Results from an Equilibrium Model of the Labor Market

F.1 Explanation of the Empirical Approach

In this appendix, we briefly describe the equilibrium model of the labor market. The full mathematical representation, formal assumptions, and equilibrium derivations are available from the authors and omitted here for brevity. The economy is composed of a large number of workers indexed by i and a large set of firms indexed by j. Each firm belongs to a market $r(j)$. Workers are heterogeneous both in preferences and productivity. Workers gain utility from after-tax wage earnings, the amenities available from firms and markets, and an idiosyncratic taste for firms. Firms differ not only in their amenities but also in terms of total factor productivity ("TFP") and technology. Workers view firms as imperfect substitutes, both due to differences in amenities offered by firms to all workers of a given quality, and due to differences in idiosyncratic tastes of workers for firms. We allow flexibility in the relationship between the amenities available from firms and markets, the market and firm-specific TFP, and the technology. Workers choose to work at the firm that maximizes utility. The idiosyncratic tastes of workers are not observed by the firm, while worker productivity is observed, and firms set profit-maximizing wages that depend on worker productivity. Under technical assumptions on the distribution of the workers' idiosyncratic preferences, we prove existence and uniqueness of the equilibrium in this economy.

Given the labor market equilibrium implied by the model, the underlying components of the model, such as TFP and amenities, are identified and can be estimated from moments of the data, the results from the BLM estimation, and the parameters of the tax function. Identification proofs and estimators are available from the authors and omitted for brevity. After estimating the underlying parameters of the model, we are able to predict from the estimated model the joint distribution of size, value added, firm effects, efficiency units of labor, and the wage bill. In Appendix Figure A.28, we provide visual evidence that the model predictions perform very well in replicating these moments, while Appendix Figure A.28e demonstrates that two estimators for the amenities offered by firms provide very similar results.[6]

First, we use the model to estimate the magnitude and sharing of rents between firms and workers, both at the firm and market level. Appendix Table A.36 presents the overall results, both when ignoring markets and when accounting for markets, while Appendix Table A.38 presents the heterogeneity in rents and rent-sharing across broad markets. Overall, we find that firm-level rents to workers are \$5,447 and firm-level rents to firms are \$5,780 in per-worker dollars, indicating that nearly 50% of rents accrue to workers. At the market-level, we find that rents to workers are \$7,331 while rents to firms are \$7,910, indicating that workers also share nearly 50% of rents at the market-level. Together, these estimates indicate that rents are large in the U.S. economy, suggesting that firms have substantial price-setting power in labor markets, but firms are not able to capture all of the rents. Table A.40 demonstrates that these results

[6] In the sample on which the model is estimated, there are 10,669,602 values of the time-varying firm premium, 61,670,459 values of worker quality, 1,953,915 values of amenity efficiency units and firm-specific TFP, 114,773 values of market-specific TFP, and 37,236,342 values of the preferences for amenities. The variances of these model components are 0.14 for amenity efficiency units, 0.04 for firm-specific TFP, 0.12 for market-specific TFP, and 0.20 for preferences for amenities.

are robust to defining an industry or region using a broader or finer concept.

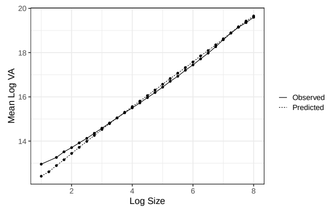
Second, we use the model to understand why firm premiums explain a small share of variance, even though the share of variance between firms is large. The results from these decompositions are reported in Table A.39. They suggest a lot of variation in amenities and productivity across firms. Interpreted in isolation, this heterogeneity predicts a large inequality contribution from firm effects. However, productive firms tend to have good amenities, which act as compensating differentials and push wages down in productive firms. As a result, firm effects explain only a few percent of the overall variation in log earnings. For example, firm effects within detailed markets explain 3.1 percent of the variation in log earnings, which is much less than predicted by the variances of firm productivity (8.6 percent) and amenities (7.1 percent).

In Appendix Figure A.29, we estimate compensating differentials directly. These use the cluster-specific estimates of ψ_j and θ_j, which are provided in Appendix Table A.37. For two randomly drawn firms, the one with worse amenities can be expected to pay an additional 18 percent in order to convince marginal workers (of average quality) to accept the job. There is, however, considerable heterogeneity in the compensating differentials according to worker quality. The upward sloping solid line shows how the expected compensating differential varies with worker quality. For high quality workers (95 percentile in the national distribution), the expected compensating differentials are as large as 30 percent. By comparison, marginal workers of low quality (5 percentile in the national distribution) require less than 10 percent additional pay to work in the firm with the unfavorable amenities.

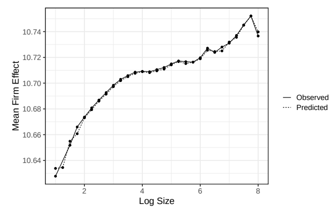
Third, we use the equilibrium model to investigate how sorting would change if we were to "shrink" the differences across firms in amenities (denoted by g_j) or in productivity interaction parameters (denoted by θ_j). Appendix Figure A.32 demonstrates how the sorting correlation and the share of log earnings variance explained by sorting varies as the amenity and productivity differences are shrunk, and Appendix Figure A.31 provides a detailed examination of how sorting responds to shrinking. We see that shrinking the g_j leads to a greater role for sorting in earnings inequality while shrinking θ_j results in a lesser role for sorting.

Fourth, in Appendix Table A.42, we compare the monopsonistic labor market to a counterfactual economy which differs in two ways. First, we eliminate the tax wedge in the first order condition by setting the tax progressivity $(1 - \lambda)$ equal to zero. Second, we remove the labor wedges in the first order conditions of the firms. Results are displayed for output, welfare, the sorting correlation, the mean labor wedge, and worker rents. They suggest the monopsonistic labor market create significant misallocation of workers to firms. Eliminating labor and tax wedges increase total welfare by 5 percent and total output by 3 percent. We also find that removing these wedges would increase the sorting of better workers to higher paying firms and lower the rents that workers earn from ongoing employment relationships.

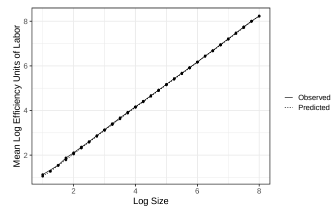
F.2 Tables and Figures



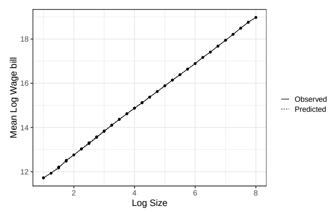
(a) Value Added



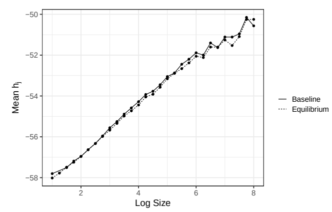
(b) Firm Effects



(c) Efficiency Units of Labor



(d) Wage Bill



(e) Fit of h_j with and without Imposing the Equilibrium Constraint

Figure A.28: Fit of the Model for Untargeted Moments

Notes: In this figure, we compare the observed and the predicted values of firm effects, value added, efficiency units of labor, and wage bill. We make this comparison separately according to the actual and predicted firm size. In subfigure e, we plot the mean of h_j across log size bins. We compare the baseline estimates of h_j from the equation for firm wage premiums, versus those estimated using the equilibrium constraint by solving the fixed-point definition of h_j as a function of amenity preferences.

	Rents and Rent-shares			
	Firm-level		**Market-level**	
Workers' Rents:				
Per-worker Dollars	5,447	(395)	7,331	(1,234)
Share of Earnings	13%	(1%)	18%	(3%)
Firms' Rents:				
Per-worker Dollars	5,780	(1,547)	7,910	(1,737)
Share of Profits	11%	(3%)	15%	(3%)
Workers' Share of Rents	49%	(4%)	48%	(3%)

Table A.36: Estimates of rents and rent sharing (national averages)

Notes: This table displays our main results on rents and rent-sharing. Column 1 presents results from the specification which imposes $\Upsilon = \gamma$, $\rho_r = 1$, and $\alpha_r = \alpha$ ("Firm only"), while columns 2-3 report results from the specification which allows Υ to differ from γ, and for ρ_r and α_r to vary across broad markets ("Accounting for Markets"). Standard errors are estimated using 40 block bootstrap draws in which the block is taken to be the market.

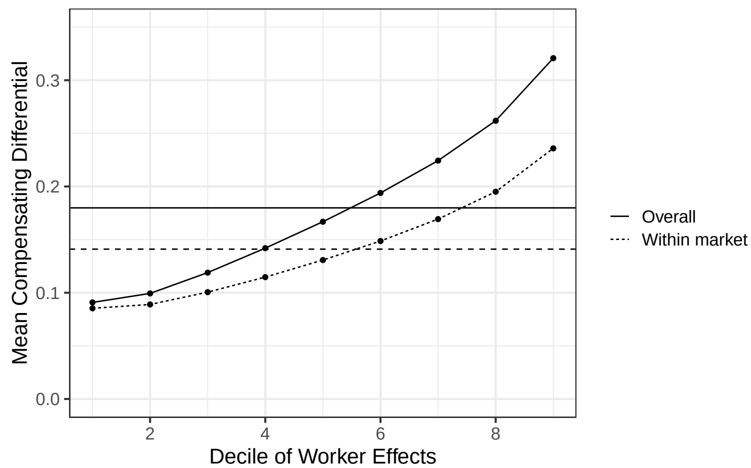


Figure A.29: Compensating differentials

Notes: In this figure, we plot mean compensating differentials overall and within market. To do so, we randomly draw a pair of firms (j, j') with probability proportional to firm size. Each j' is drawn from the full set of firms when estimating overall compensating differentials and from the set of firms in the same market as j when estimating within market compensating differentials. Then, we estimate the compensating differential between j and j' for a worker of given quality $x_i = x$ by $\psi_j + x\theta_j - \psi_{j'} - x\theta_{j'}$. This figure plots the mean absolute value of the compensating differentials across deciles of the x_i distribution, where the horizontal lines denote means across the distribution of x_i.

	No Time-varying Effects							With Time-varying Effects						
					$\mathbb{E}[w\|x,k]$ by x quantile							$\mathbb{E}[w\|x,k]$ by x quantile		
Cluster k	N_k	$\mathbb{E}\log W$	ψ_k	θ_k	20th	50th	80th	N_k	$\mathbb{E}\log W$	ψ_k	θ_k	20th	50th	80th
1	11.16	10.10	0.00	1.00	-0.68	-0.43	-0.18	11.06	10.10	0.00	1.00	-0.67	-0.43	-0.18
2	24.88	10.28	0.17	1.12	-0.59	-0.31	-0.03	24.71	10.28	0.17	1.13	-0.58	-0.31	-0.03
3	32.82	10.44	0.38	1.39	-0.57	-0.22	0.13	32.70	10.43	0.39	1.41	-0.56	-0.21	0.14
4	32.59	10.56	0.45	1.42	-0.52	-0.16	0.19	32.67	10.56	0.47	1.45	-0.50	-0.15	0.21
5	23.19	10.70	0.51	1.43	-0.46	-0.10	0.25	23.41	10.70	0.54	1.47	-0.45	-0.09	0.27
6	21.21	10.75	0.65	1.81	-0.59	-0.14	0.32	21.05	10.75	0.68	1.88	-0.58	-0.12	0.34
7	28.39	10.87	0.67	1.69	-0.48	-0.06	0.36	28.45	10.87	0.69	1.74	-0.47	-0.05	0.38
8	24.26	11.07	0.77	1.89	-0.51	-0.04	0.43	24.38	11.07	0.80	1.94	-0.51	-0.03	0.44
9	18.28	11.30	0.89	2.02	-0.49	0.01	0.52	18.49	11.30	0.91	2.08	-0.48	0.03	0.54
10	9.02	11.63	1.01	2.24	-0.52	0.04	0.60	8.88	11.64	1.03	2.31	-0.52	0.05	0.62

Table A.37: Parameter Estimates with Firm-Worker Interactions

Notes: In this table, we describe the estimated parameters and wage predictions from the BLM specification with firm-worker interactions. $\log W$ refers to the mean of (non-residual) observed log earnings within a group k. The model prediction for (residual) log earnings is given by $\mathbb{E}\left[w|x = x_q, k = k\right] = \psi_k + \theta_k \cdot x_q$, for each group $k = 1, 2, ..., 10$ and considering various quantiles x_q in the distribution of x. N_k refers to the number of worker-years (in millions) observed in the cluster during the 2001-2008 time interval.

	Goods				Services			
	Midwest	Northeast	South	West	Midwest	Northeast	South	West
Panel A.	**Model Parameters**							
Idyosinctratic taste parameter (β^{-1})	0.200							
	(0.044)							
Taste correlation parameter (ρ)	0.844	0.694	0.719	0.924	0.649	0.563	0.744	0.619
	(0.179)	(0.153)	(0.160)	(0.182)	(0.141)	(0.109)	(0.246)	(0.117)
Returns to scale ($1-\alpha$)	0.746	0.764	0.863	0.949	0.753	0.740	0.814	0.752
	(0.016)	(0.013)	(0.017)	(0.019)	(0.013)	(0.015)	(0.036)	(0.015)
Panel B.	**Firm-level Rents and Rent Shares**							
Workers' Rents:								
Per-worker Dollars	6,802	6,681	5,737	8,906	4,234	4,847	5,009	4,805
	(770)	(723)	(720)	(867)	(502)	(803)	(1,295)	(684)
Share of Earnings	16%	13%	14%	17%	12%	11%	14%	12%
	(2%)	(1%)	(2%)	(2%)	(1%)	(2%)	(4%)	(2%)
Firms' Rents:								
Per-worker Dollars	4,041	4,198	7,465	20,069	3,531	3,097	6,915	3,018
	(1,243)	(1,130)	(2,681)	(6,323)	(1,004)	(1,305)	(5,650)	(1,060)
Share of Profits	8%	7%	17%	52%	6%	5%	12%	6%
	(3%)	(2%)	(6%)	(16%)	(2%)	(2%)	(10%)	(2%)
Workers' Share of Rents	63%	61%	43%	31%	55%	61%	42%	61%
	(4%)	(4%)	(5%)	(4%)	(4%)	(5%)	(9%)	(5%)
Panel C.	**Market-level Rents and Rent Shares**							
Workers' Rents:								
Per-worker Dollars	7,837	9,102	7,572	9,506	6,115	7,935	6,422	7,230
	(1,319)	(1,532)	(1,274)	(1,600)	(1,029)	(1,335)	(1,081)	(1,217)
Share of Earnings	18%	18%	18%	18%	18%	18%	18%	18%
	(3%)	(3%)	(3%)	(3%)	(3%)	(3%)	(3%)	(3%)
Firms' Rents:								
Per-worker Dollars	4,940	6,311	10,000	20,846	5,734	5,897	9,363	5,153
	(1,140)	(1,350)	(2,267)	(5,787)	(1,351)	(1,786)	(4,218)	(1,433)
Share of Profits	10%	11%	23%	54%	10%	9%	16%	10%
	(2%)	(2%)	(5%)	(15%)	(2%)	(3%)	(7%)	(3%)
Workers' Share of Rents	61%	59%	43%	31%	52%	57%	41%	58%
	(3%)	(3%)	(4%)	(5%)	(3%)	(4%)	(8%)	(4%)

Table A.38: Broad Market Heterogeneity in Model Parameters and Rent Sharing Estimates

Notes: This table displays our heterogeneity in the estimated model parameters and rents and rent-sharing. These results correspond to the specification which allows Υ to differ from γ, and for ρ_r and α_r to vary across broad markets. Standard errors are estimated using 40 block bootstrap draws in which the block is taken to be the market.

	Between Broad Markets	**Within Broad Markets**	
		Between Detailed Markets	**Within Detailed Markets**
	Preferred Specification		
Total	0.4%	2.0%	3.1%
Decomposition:			
Amenity Differences	16.0%	7.8%	7.1%
TFP Differences	15.5%	11.9%	8.6%
Amenity-TFP Covariance	-31.1%	-17.7%	-12.6%
	Log-additive Fixed Effects Specification		
Total	0.6%	2.8%	6.6%
Decomposition:			
Amenity Differences	15.7%	6.5%	7.2%
TFP Differences	14.6%	13.2%	10.0%
Amenity-TFP Covariance	-29.8%	-16.9%	-10.5%

Table A.39: Decomposition of the Variation in Firm Premiums

Notes: This table displays our estimates of the decomposition of time-varying firm premium variation in three levels: variation between broad markets, between detailed markets (within broad markets), and between firms (within detailed markets). Broad markets are defined as the combination of sector times region, and detailed markets are defined as the combination of industry times commuting zone. We decompose the variation in time-varying firm premiums into the contributions from amenity differences, TFP differences, and the covariance between amenity and TFP differences. All variances are expressed as shares of log earnings variance. Results are presented both for the preferred BLM specification and the AKM specification. For reference, the variance of g_j is about 0.20 while the variances of $\tilde{a}_{jt}$ and $\bar{a}_{rt}$ are about 0.14 and 0.12, respectively.

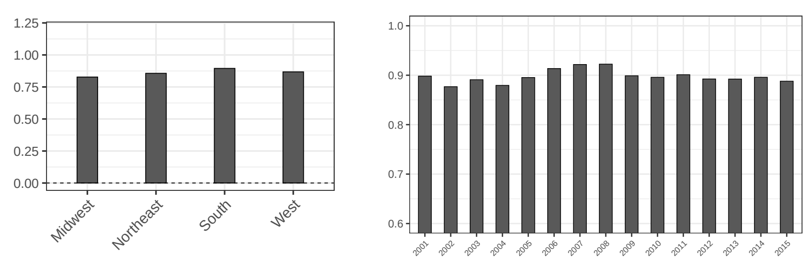
	Market Count (in 1,000)		Passthrough Rate		Average of the Model Parameters			Workers' Share of Rents Firm-level	Market-level
	Workers	Firms	Market	Firm	β	$1-\rho_r^2$	$1-\alpha_r$	$\frac{R^w}{R^w+R^f}$	$\frac{R^{wm}}{R^{wm}+R^{fm}}$
Baseline (NAICS 2-digit, commuting zone)	1.90	0.17	0.18	0.13	4.99	0.51	0.79	0.52	0.50
Shutdown broad market heterogeneity ($\rho_r = \overline{\rho}, \alpha_r = \overline{\alpha}$)	1.97	0.17	0.18	0.13	5.06	0.48	0.79	0.52	0.51
Alternative detailed markets:									
Finer geography (county)	0.54	0.05	0.19	0.14	4.61	0.54	0.79	0.51	0.49
Finer industry (NAICS 3-digit)	0.65	0.06	0.19	0.13	4.60	0.59	0.79	0.52	0.50
Coarser geography (state)	25.44	2.23	0.18	0.13	5.00	0.52	0.79	0.53	0.50
Coarser industry (NAICS supersector)	4.42	0.39	0.20	0.13	4.28	0.66	0.79	0.53	0.51

Table A.40: Robustness of the Model Parameters and Rent Sharing Estimates to Alternative Market Definitions

Notes: This table displays robustness of the estimated model parameters and rents to alternative definitions of detailed markets.

Panel A. Technology and Product Demand Parameters			
	Baseline Estimates using Over-identified GMM		
	Parameters	**Data**	
Private demand parameter	$1-\epsilon$	0.863	(0.015)
Composite labor scale parameter	ρ	1.089	(0.017)
Returns to labor parameter	β_L	0.499	(0.192)
	Alternative Estimates using Exactly-identified OLS		
	Parameters	**Data**	
Diminishing returns to output (eq 22)	$1-\epsilon$	0.863	(0.008)
Optimal intermediate inputs to employees (eq 25)	ρ	1.057	(0.015)
Labor to value added ratio (eq 26)	β_L	0.514	(0.209)
Panel B. Remaining Parameters for Price, Scale, and TFP			
	Parameter and Identifying Moments	**Data**	
Returns to capital	$\beta_K = (\rho - \beta_L)/(1+\theta)$	0.474	(0.161)
Scale of optimal log wage	$\mathbb{E}[u_{jt}] = \mathbb{E}[b_{jt}] - (1+\theta)\mathbb{E}[\ell_{jt}]$	10.075	(0.000)
Scale term for intermediates	$\log \frac{\beta_M}{p_M} = \rho\mathbb{E}[\ell_{jt}] - \mathbb{E}[x_{jt}]$	-11.722	(0.047)
Scale of log output price	$\log p_H = \mathbb{E}[r_{jt} - (1-\epsilon)(\log \frac{\beta_M}{p_M} + x_{jt})\vert D=0]$	12.801	(0.053)
Interquartile range of log TFP	$\mathrm{IQR}(\phi_{jt}) = \mathrm{IQR}(x_{jt} - \rho\ell_{jt})$	0.918	(0.001)

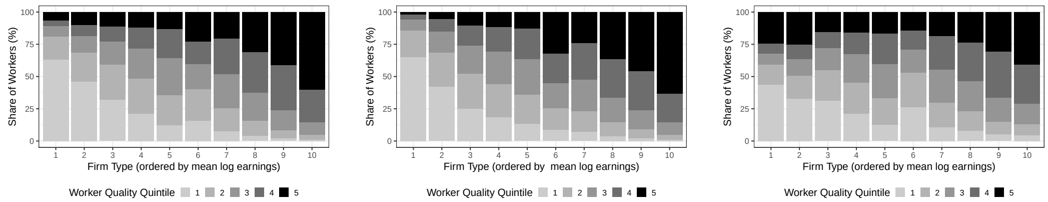
Table A.41: Firm Technology and Product Demand Parameters



(a) Estimate of $1-\epsilon$ by Broad Region (b) Interquartile Range in TFP by Year

Figure A.30: Heterogeneity in Model Parameters

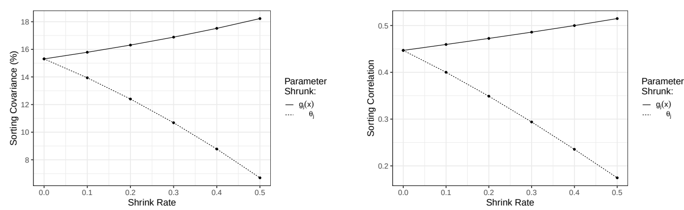
Notes: In this figure, we demonstrate the heterogeneity in $1-\epsilon$ across broad regions, as well as the heterogeneity in the interquartile range of the TFP estimates across years.



(a) Baseline Equilibrium (b) Shrink $g_j(x)$ (c) Shrink θ_j

Figure A.31: Actual and counterfactual composition of the workforce by firm types

Notes: In this figure, we reduce the heterogeneity across firms in amenities or production complementarities by replacing either $g_j(x)$ with $(1-s)\, g_j(x) + s\bar{g}_j$ or θ_j with $(1-s)\,\theta_j + s\bar{\theta}$, where $\bar{g}_j = \mathbb{E}_x\left[g_j(x)\right]$, $\bar{\theta} = \mathbb{E}\left[\theta_j\right]$. Here, $s \in [0, 1]$ is the shrink rate with $s = 0$ corresponding to the baseline model. We report the quality of the workforce by firm type in the baseline economy with $s = 0$ (subfigure a) and the counterfactual economies with $s = \frac{1}{2}$ for either amenities (subfigure b) or production complementarities (subfigure c).



(a) Sorting (b) Sorting Correlation

Figure A.32: Worker sorting with counterfactual values of $g_j(x)$ and θ_j

Notes: In this figure, we reduce the heterogeneity across firms in amenities or production complementarities by replacing either $g_j(x)$ with $(1-s)\, g_j(x) + s\bar{g}_j$ or θ_j with $(1-s)\,\theta_j + s\bar{\theta}$, where $\bar{g}_j = \mathbb{E}_x\left[g_j(x)\right]$, $\bar{\theta} = \mathbb{E}\left[\theta_j\right]$. Here, $s \in [0, 1]$ is the shrink rate with $s = 0$ corresponding to the baseline model. We report the share of log earnings variance explained by sorting (subfigure a) and the sorting correlation (subfigure b).

		(1) Monopsonistic Labor Market	(2) No Labor or Tax Wedges	Difference between (1) and (2)
Log of Expected Output	$\log \mathbb{E}[Y_{jt}]$	11.38	11.41	0.03
Total Welfare (log dollars)		12.16	12.21	0.05
Sorting Correlation	$Cor(\psi_{jt}, x_i)$	0.44	0.47	0.03
Labor Wedges	$1 + \frac{\rho_r}{\beta\lambda}$	1.15	1.00	-0.15
Worker Rents (as share of earnings):				
Firm-level	$\frac{\rho_r}{\rho_r+\beta\lambda}$	13.3%	12.3%	-1.0%
Market-level	$\frac{1}{1+\beta\lambda}$	18.0%	16.7%	-1.3%

Table A.42: Consequences for Worker Allocation and Outcomes of Eliminating Tax and Labor Wedges

Notes: This table compares the monopsonistic labor market to a counterfactual economy which differs in two ways. First, we eliminate the tax wedge in the first order condition by setting the tax progressivity $(1-\lambda)$ equal to zero. Second, we remove the labor wedges in the first order conditions of the firms by setting τ_r equal to the labor wedge $1 + \frac{\rho_r}{\lambda\beta}$ in each market r. After changing these parameters of the model, we solve for the new equilibrium allocation and outcomes, including wages, output and welfare. Results are displayed for output, welfare, the sorting correlation, the mean labor wedge, and worker rents.