

NBER WORKING PAPER SERIES

DO FIRM EFFECTS DRIFT? EVIDENCE FROM
WASHINGTON ADMINISTRATIVE DATA

Marta Lachowska
Alexandre Mas
Raffaele D. Saggio
Stephen A. Woodbury

Working Paper 26653
<http://www.nber.org/papers/w26653>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020, Revised October 2021

We thank David Card, Patrick Kline, Christian Moser, Martha Stinson, Mikkel Sølvsten, Ian Schmutte, Jenna Stearns and Lars Vilhuber for helpful comments. We are grateful to the Employment Security Department (ESD) of Washington State for allowing access to the Washington wage records, and especially to Jeff Robinson of ESD, whose help was essential to understanding the data. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Marta Lachowska, Alexandre Mas, Raffaele D. Saggio, and Stephen A. Woodbury. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Do Firm Effects Drift? Evidence from Washington Administrative Data
Marta Lachowska, Alexandre Mas, Raffaele D. Saggio, and Stephen A. Woodbury
NBER Working Paper No. 26653
January 2020, Revised October 2021
JEL No. J0, J3

ABSTRACT

We study the time-series properties of firm effects in the two-way fixed effects model popularized by Abowd, Kramarz, and Margolis (1999) (AKM) using two approaches. The first—the rolling AKM approach (R-AKM)—estimates AKM models separately for successive two-year intervals. The second—the time-varying AKM approach (TV-AKM)—is an extension of the original AKM model that allows for unrestricted interactions of year and firm indicators. We apply to both approaches the leave-out methodology of Kline, Saggio and Sølvssten (2020) to correct for biases in the estimated variance components. Using administrative wage records from Washington State, we find, first, that firm effects for hourly wage rates are highly persistent with an autocorrelation coefficient between firm effects in 2002 and 2014 of 0.74. Second, the R-AKM approach reveals cyclicalities in firm effects and worker-firm sorting. During the Great Recession the variability in firm effects increased, while the degree of worker-firm sorting decreased. Third, misspecification of standard AKM models resulting from restricting firm effects to be fixed over time appears to be minimal.

Marta Lachowska
W.E. Upjohn Institute for
Employment Research
300 S. Westnedge Avenue
Kalamazoo, MI 49007
marta@upjohn.org

Alexandre Mas
Industrial Relations Section
Simpson International Building
Princeton University
Princeton, NJ 08544
and NBER
amas@princeton.edu

Raffaele D. Saggio
Department of Economics
University of British Columbia
6000 Iona Drive
Vancouver, BC V6T 1L4
and NBER
rsaggio@mail.ubc.ca

Stephen A. Woodbury
Department of Economics
Michigan State University
486 West Circle Drive, Room 110
East Lansing, MI 48824
and W.E. Upjohn Institute for
Employment Research
woodbur2@msu.edu

1 Introduction

The variance decomposition method proposed by Abowd, Kramarz, and Margolis (AKM, 1999) has been the workhorse of a large and growing literature examining earnings differentials and structural change in labor markets. By decomposing earnings variability into components attributable to workers, firms, and the sorting of workers to firms, it has produced a wealth of insights regarding labor market inequality (Card, Heining and Kline, 2013; Song, Price, Guvenen, Bloom and von Wachter, 2019), gender wage differences (Card, Cardoso and Kline, 2015), compensating differentials for firm characteristics (Sorkin, 2018), the influence of outsourcing on earnings (Goldschmidt and Schmieder, 2017), and the sources of displaced workers’ earnings losses (Schmieder, von Wachter and Heining, 2018; Lachowska, Mas and Woodbury, 2020).

The standard AKM model imposes the assumption that firm effects—the contribution of a given firm’s pay policies to workers’ wage—are time invariant. This assumption allows researchers to pool multiple years of data to estimate the model parameters. Pooling is helpful because it is known that the simple “plug-in” estimator commonly used to obtain the AKM variance components is biased and inconsistent due to sampling error in the estimated worker and firm effects, as originally noted by Krueger and Summers (1988) and Abowd et al. (2004). This bias is particularly severe in datasets with few worker transitions—Andrews, Gill, Schank and Upward (2008) refer to this as limited mobility bias. The assumption of time-invariant firm effects justifies pooling multiple years of data to increase the number of observed worker transitions, reducing sampling error and alleviating the bias in estimating variance components.

However, the literature on rent sharing provides evidence that firm pay policies may change over time. For example, Van Reenen (1996) and Kline et al. (2019) document that pay is related to time-varying patent activity, and a growing literature has estimated the co-movement of firm performance and compensation (Guiso, Pistaferri and Schivardi, 2005; Card, Devicienti and Maida, 2014; Card, Cardoso, Heining and Kline, 2018). By pooling many time periods and assuming that firm pay policies are time invariant, we risk understating the true variability of firm pay policies.

In this paper, we extend the original AKM model to allow for time-varying firm heterogeneity

while controlling for biases in the resulting variance components due to limited mobility. We propose two extensions. The first—the rolling AKM approach (R-AKM)—estimates AKM models separately for successive two-year time intervals. The second—the time-varying AKM approach (TV-AKM)—is based on an extension of the original AKM model where we allow for unrestricted interactions of year and firm indicators. We estimate both models using Washington State administrative wage records from 2002 through 2014, and then apply the recent leave-out bias correction of Kline, Saggio and Sølvssten (2020) (KSS) to correct for limited mobility bias in the resulting variance components.¹

The KSS correction, applied to both the R-AKM and TV-AKM models, produces unbiased parameter estimates that we believe are of interest to both labor and macro-labor economists. For example, the R-AKM estimates can be used to examine the cyclicity of firm wage policies (whether firm heterogeneity in wages increases during recessions) and whether sorting of workers to firms varies over the business cycle. Similarly, the TV-AKM model is well-suited to understanding the time-series properties of wage policies (whether firm effects are strongly correlated over time).

The main findings are as follows. First, firm wage effects show a remarkable degree of stability. Specifically, in a balanced panel of firms observed in every year during 2002–2014, the autocorrelation coefficient between firm effects in 2002 and 2014 is 0.74. We find that firm effects are well approximated by a persistent AR(1) process. The stability of firm effects provides evidence that, by and large, firm effects represent permanent differences in firm compensation policies.

Second, we find that the standard AKM and TV-AKM estimates of variance components are quite similar. Specifically, we find the KSS bias-corrected variance of time-varying firm effects explain 12 percent of the variance of wages while time-invariant firm effects explain about 11 percent. We also find that, when pooling multiple years of data, KSS bias-corrected estimates of the variance components are very close to those obtained using a simple plug-in approach.

Third, we find that two of the key components of wage variation showed cyclicity over the

¹In contemporaneous work, Engbom and Moser (2020) also analyze variance decompositions based on a TV-AKM model and correlate the estimated firm-year effects with measures of firm productivity using data from Sweden. Rather than using the KSS correction, they address the limited mobility bias by presenting estimates for alternative firm-size thresholds.

years we examine. Specifically, the dispersion of firm effects fell in the years before the Great Recession (2002–2007), increased during the Great Recession, then decreased dramatically in the post-recession years (after 2010). In contrast, the covariance between worker and firm effects was pro-cyclical, growing in the years before the Great Recession, decreasing significantly during the recession, and increasing dramatically in the post-recession years. These cyclical and opposing patterns in the dispersion of firm effects and in the sorting of workers to firms appear to be novel findings. We also examine how firm effects, worker effects, and worker-firm sorting estimated separately in each interval contribute to the observed increases in wage inequality in Washington during 2002–2014. Bias-corrected estimates suggest that the increase in wage inequality can be attributed almost entirely to increases in the variation of worker-specific effects and increased sorting of “good” workers to “good” firms, similar to the findings in Song, Price, Guvenen, Bloom and von Wachter (2019).

These main results are further scrutinized in two ways. First, we highlight key differences between the R-AKM and TV-AKM model. The identification of firm effects in R-AKM hinges solely on the transitions made by workers in that interval. The set of identified firm effects in R-AKM is therefore potentially much smaller compared to the set in the TV-AKM, as the latter exploits information on both stayers and workers’ transitions over 13 years. This difference in the identified sets of firms matters because we tend to find less variability in firm effects using R-AKM when compared to TV-AKM.

Second, we elaborate on the finding that the KSS correction is more important in shorter panels than in longer panels by the reporting the magnitude of the KSS correction for panels of various length. We argue that a better statistic to assess the importance of correcting for limited mobility bias KSS correction in an AKM model is the ratio of the number of movers to the number of firms, rather than the length of the panel per se.

The analysis concludes with three extensions of the main analysis. First, two simple applications provide evidence about the degree of misspecification stemming from restricting firm effects to be fixed over time. Regressing TV-AKM firm wage effects on time-invariant AKM firm wage ef-

fects results in a projection slope close to unity. Also, regressing separation rates on time-invariant AKM firm wage effects results in virtually the same estimated relationship as regressing separation rates on TV-AKM firm wage effects. The conclusion from both exercises is that misspecification from assuming time invariance of firm effects is likely a second-order concern. Second, we show that the firm effects of new firms tend to be less stable than those of older firms: for new firms, the autocorrelation coefficient between firm effects for 2003 and 2014 is 0.46 (compared to a comparable autocorrelation coefficient for all firms' fixed effects of 0.79). Third, we show that worker effects are nearly as persistent as firm effects. Based on estimates from the R-AKM model, the autocorrelation coefficient between worker effects estimated in 2002–2003 and 2014–2014 is 0.70.

The paper is organized as follows. Section 2 develops the econometric framework, reviewing both the KSS bias-correction estimator and the methods of estimating time-varying firm effects. Section 3 describes the data, connectivity in alternative samples, and the growth of wage inequality in Washington during 2002–2014. Section 4 discusses the decomposition of wage variability and the cyclicity of these components, based on fitting AKM models on two-year adjacent intervals (the R-AKM model). Section 5 describes the time-series properties of the firm-year effects obtained after fitting an AKM model with time-varying firm fixed effects (the TV-AKM model). Section 6 examines differences between the sets of firms used to estimate the R-AKM and TV-AKM models and also examines the importance of firm entry and exit. Section 7 discusses the distinction between the length of a panel and degree of mobility in alleviating limited mobility bias. Section 8 investigates three further questions: the degree of misspecification implied by imposing the assumption of fixed firm effects, the stability of new firms' fixed effects, and whether worker effects (as opposed to firm effects) drift. Section 9 offers some concluding remarks. Appendices offer a more detailed description of the data and computational aspects of the KSS correction.

2 Econometric Framework

The starting point of our discussion is the standard two-way fixed effects model popularized by AKM. In this model, the log hourly wage of worker g at year t , y_{gt} , is decomposed into the sum of

a worker component, α_g , a firm component, ψ_j , and an error component ε_{gt} :

$$y_{gt} = \alpha_g + \psi_{j(g,t)} + w'_{gt} \delta + \varepsilon_{gt}. \quad (1)$$

The function $j(\cdot, \cdot) : \{1, \dots, N\} \times \{1, \dots, \max_g T_g\} \rightarrow \{1, \dots, J\}$ allocates each of $n = \sum_{g=1}^N T_g$ worker-year observations to one of J firms, where T_g denotes the total number of years in which we observe worker g . In equation (1), α_g is a worker effect that captures a combination of time-invariant skills and other factors of a given worker that are rewarded equally across different firms. The term ψ_j represents a firm-specific relative pay premium that is paid equally by firm j to all its employees. The term w_{gt} capture some time-varying controls.² Finally, ε_{gt} represents an unobserved time-varying error that captures random match effects, shocks to human capital, and other unobserved factors.

The AKM model is a useful tool to assess the influence of firms in setting wages. For simplicity, consider the case where $\delta = 0$ in (1). Then, according to equation (1), the variance of log wages can be decomposed as

$$\text{var}(y_{gt}) = \text{var}(\alpha_g) + \text{var}(\psi_{j(g,t)}) + 2\text{cov}(\psi_{j(g,t)}, \alpha_g) + \text{var}(\varepsilon_{gt}). \quad (2)$$

This decomposition highlights that, if firms have sufficient latitude in setting wages, wage inequality will be affected through two terms: $\text{var}(\psi_{j(g,t)})$ and $\text{cov}(\psi_{j(g,t)}, \alpha_g)$, where the latter captures the extent to which workers with a higher fixed wage component tend to sort into firms paying higher wage premiums. The growing availability of large administrative datasets and improvements in computational methods has allowed economists to provide new evidence on the contribution of firms and worker-firm sorting to wage inequality—e.g., Card, Heining and Kline (2013) for Germany; Song, Price, Guvenen, Bloom and von Wachter (2019) for the US.

Two econometric challenges arise when interpreting results from these recent studies. First, most of the available evidence on the importance of firms and worker-firm sorting is based on a simple “plug-in” approach, where each variance component in (2) is calculated as the variance

²In our application, these are limited to year effects.

of OLS estimates of (α, ψ) from equation (1). However, as initially pointed out by Krueger and Summers (1988) and Abowd et al. (2004), sampling error in the estimates of (α, ψ) will impart bias in the estimated variance components, a phenomenon often referred to as “limited mobility bias.” These biases can be particularly severe when the number of workers transitioning between different employers is relatively low compared to the overall dimensionality of the model (Andrews, Gill, Schank and Upward, 2008; Kline, Saggio and Sølvesten, 2020).

Second, the original AKM model assumes the firm effects $\psi_{j(g,t)}$ to be time invariant. This assumption could be problematic if, in an attempt to minimize limited mobility bias, researchers end up estimating the original AKM model by pooling, say, 10 or more years of data (Goldschmidt and Schmieder, 2017; Bana, Bedard, Rossin-Slater and Stearns, 2018; Lachowska, Mas and Woodbury, 2020). However, it is unclear why firm wage policies should remain fixed over such long horizons. In fact, recent evidence suggests firm wage policies are sensitive to within-firm idiosyncratic shocks (Kline, Petkova, Williams and Zidar, 2019; Garin, Silvério et al., 2019). A more realistic representation of the wage-setting process is therefore one where firm effects are allowed to vary over time.³

These two challenges are intimately related. Economists are interested in estimating flexible models that allow firm effects to vary over time. But limited mobility bias has thus far constrained these attempts. This paper shows that both challenges can be addressed using the KSS bias-correction procedure, which allows computation of unbiased estimates of variance components in *any* high-dimensional linear model, including a model with time-varying firm effects.

We next present a brief overview of the KSS leave-out methodology when applied to an AKM model. We then introduce two extensions of the AKM model: the Rolling AKM (R-AKM) and the Time-varying AKM (TV-AKM), both of which allow firm effects to vary over time.

³For instance, Lamadon, Mogstad and Setzler (2019) develop a model where standard AKM firm effects vary over time whenever pass-through rent sharing elasticities differ from zero and a given firm experiences an idiosyncratic change to its value added.

2.1 The KSS Correction in the Standard AKM Model

Suppose we are interested in estimating the firm-size weighted variability of firm effects

$$\sigma_{\bar{\psi}}^2 \equiv \text{var}(\psi_{j(g,t)}) = \frac{1}{n} \sum_{g=1}^N \sum_{t=1}^{T_g} (\psi_{j(g,t)} - \bar{\psi})^2 \quad (3)$$

where $\bar{\psi} = \frac{1}{n} \sum_{g=1}^N \sum_{t=1}^{T_g} \psi_{j(g,t)}$. Without loss of generality, assume $\bar{\psi} = 0$, in which case we can rewrite $\sigma_{\bar{\psi}}^2$ simply as

$$\sigma_{\bar{\psi}}^2 = \sum_{j=1}^J s_j \psi_j^2 \quad (4)$$

where s_j gives the employment share of firm j , $s_j = \frac{1}{n} \sum_{g=1}^N \sum_{t=1}^{T_g} \mathbf{1}\{j(g,t) = j\}$. The “plug-in” approach estimates $\sigma_{\bar{\psi}}^2$ by replacing ψ_j with its corresponding OLS estimate $\hat{\psi}_j$ obtained from fitting equation (1)

$$\widetilde{\sigma}_{\bar{\psi}}^2 = \sum_{j=1}^J s_j \hat{\psi}_j^2. \quad (5)$$

We next discuss the bias associated with $\widetilde{\sigma}_{\bar{\psi}}^2$ also known as “limited mobility bias”.

2.1.1 Limited Mobility Bias

The estimated firm effect, $\hat{\psi}_j$, represents a noisy estimate of the true firm effect, ψ_j . The presence of noise in $\hat{\psi}_j$ is not an issue when one is interested in ψ_j as the OLS estimator $\hat{\psi}_j$ is assumed to be unbiased, i.e., $E[\hat{\psi}_j] = \psi_j$.

However, the estimation error in $\hat{\psi}_j$ is going to lead to biases if one is interested in estimating ψ_j^2 using its “plug-in” analogue $\hat{\psi}_j^2$ since

$$E[\hat{\psi}_j^2] = E[(\hat{\psi}_j - \psi_j + \psi_j)^2] = \psi^2 + \underbrace{V[\hat{\psi}_j]}_{\text{Bias}}, \quad (6)$$

where $V[\hat{\psi}_j]$ is the squared standard error of $\hat{\psi}_j$. Intuitively, when we take the square of $\hat{\psi}_j$ we are not only squaring its signal, ψ_j , but also the estimation error in each $\hat{\psi}_j$. The latter is going

to introduce a bias when estimating ψ_j^2 . The same logic applies when analyzing the bias of the plug-in estimator of the variance of firm effects because

$$E[\widetilde{\sigma}_{\psi}^2] = \sigma_{\psi}^2 + \underbrace{\sum_{j=1}^J s_j V[\hat{\psi}_j]}_{\text{Bias}}. \quad (7)$$

2.1.2 Why Standard Methods Fail to Correct for Limited Mobility Bias

It is useful to rewrite model (1) as follows

$$y_i = d_i' \alpha + f_i' \psi + \varepsilon_i = x_i' \beta + \varepsilon_i \quad (8)$$

where i indexes a particular worker-year observation (g, t) ; d_i and f_i denote worker and firm identifiers respectively and we have $x_i = (d_i', f_i)'$, $\beta = (\alpha', \psi)'$ with $\alpha = (\alpha_1, \dots, \alpha_N)$ and $\psi = (\psi_1, \dots, \psi_J)$.

Equation (7) shows that to correct the bias in the plug-in estimate of the variance of firm effects, we need an estimate of the (squared) standard error of each firm effect, $V[\hat{\psi}_j]$. Similarly, if we are interested on the variance of worker effects, then we need the standard error on each of the worker effects, $V[\hat{\alpha}_i]$. If we are interested in the covariance of worker and firm effects, then we need the covariances in sampling errors between each $\hat{\alpha}_i$ and $\hat{\psi}_{j(i,t)}$. In other words, all that is required to correct for the limited mobility bias is an estimate of the sampling variability of the OLS coefficient vector $\hat{\beta}$, that is

$$V[\hat{\beta}] = \left(\sum_{i=1}^n x_i x_i' \right)^{-1} \sum_{i=1}^n \sigma_i^2 x_i x_i' \left(\sum_{i=1}^n x_i x_i' \right)^{-1}, \quad (9)$$

where $\sigma_i^2 = \text{Var}(\varepsilon_i)$.

One might be tempted to provide an estimate of $V[\hat{\beta}]$ using heteroskedasticity *consistent* (“HC”) or robust standard errors. Standard White (1980) HC standard-errors are calculated using a plug-in estimate of σ_i^2 based on

$$\tilde{\sigma}_i^2 = (y_i - x_i' \hat{\beta})^2, \quad (10)$$

and the resulting HC-based estimate of $V[\hat{\beta}]$ is then given by

$$\tilde{V}[\hat{\beta}] = \left(\sum_{i=1}^n x_i x_i' \right)^{-1} \sum_{i=1}^n \tilde{\sigma}_i^2 x_i x_i' \left(\sum_{i=1}^n x_i x_i' \right)^{-1}. \quad (11)$$

However, HC standard errors based on $\tilde{\sigma}_i^2$ are downward biased (MacKinnon and White, 1985). Moreover, HC standard errors are inconsistent in any high-dimensional model where the number of parameters grows in proportion to the sample size (Cattaneo, Jansson and Newey, 2018). Such “many regressor” asymptotics are common in the AKM models where, often there are fewer than 5 worker moves, on average, per firm. Hence, traditional HC standard errors cannot be used in this context.

2.1.3 The Leave-Out Correction

KSS derives a novel heteroskedasticity-*unbiased* (HU) estimate of the standard error of any coefficient obtained from a linear regression model (including high-dimensional models, such as AKM). The KSS HU standard error estimate is based on a leave-out estimate of σ_i^2 :

$$\hat{\sigma}_i^2 = y_i(y_i - x_i' \hat{\beta}_{-i}), \quad (12)$$

where $\hat{\beta}_{-i}$ is the OLS estimate of β from equation (2) when observation i is left out.

KSS then replaces σ_i^2 in $V[\hat{\beta}]$ with its unbiased estimate $\hat{\sigma}_i^2$ to derive an HU estimate of $V[\hat{\beta}]$:

$$\hat{V}[\hat{\beta}] = \left(\sum_{i=1}^n x_i x_i' \right)^{-1} \sum_{i=1}^n \hat{\sigma}_i^2 x_i x_i' \left(\sum_{i=1}^n x_i x_i' \right)^{-1}. \quad (13)$$

Returning to the variance of firm effects, we can extract the corresponding squared standard error of each firm effect, $\hat{V}[\hat{\psi}_j]$, from $\hat{V}[\hat{\beta}]$. We can then use it to bias-correct the corresponding estimate of the variance of the firm effects as follows:

$$\widehat{\sigma}_{\psi}^2 = \widetilde{\sigma}_{\psi}^2 - \sum_{j=1}^J s_j V[\widehat{\psi}_j]. \quad (14)$$

In this paper, we use KSS to correct all the variance components associated with AKM, as well as our proposed extensions of AKM that allows for time-varying heterogeneity. Appendix B.3 provides details on implementation.

2.2 Time-Varying Firm Heterogeneity

The previous section explained how to apply the KSS correction to a standard AKM model. We now turn to the extensions that allow for time-varying firm heterogeneity in AKM.

2.2.1 Rolling AKM (R-AKM)

A straightforward approach that permits firm effects to vary over time is to estimate the standard AKM model separately for different time intervals. This was the approach taken by Card, Heining and Kline (2013) who estimated separate AKM models across approximately 6-year intervals. The KSS correction permits us to estimate AKM models at a much higher interval frequency without worrying about potential issues caused by limited mobility biases.

Specifically, we can fit equation (1) to successive overlapping two-year ($T = 2$) intervals (2002–2003, 2003–2004, ..., 2013–2014) and correct the associated variance decompositions for each time interval using the KSS approach. We label this approach rolling AKM (R-AKM). The key advantage of R-AKM is its simplicity and transparency. We view it as a particular effective tool to summarize the cyclical pattern of the AKM variance components.

It is useful to note that the R-AKM model allows *both* firm and worker effects to vary between two-year intervals. This is because the R-AKM model is a standard AKM model estimated separately over a sequence of very short intervals. In this model, a worker who is observed in interval t and in $t + j$ is not constrained to have the same worker effect in the two intervals. Therefore, we can estimate the degree of autocorrelation between a worker’s effects estimated j intervals apart.

Analogously, firm effects are also allowed to vary across intervals. This effectively means that

firm effects in interval t are estimated solely by the workers' transitions observed in interval t . We cannot identify a firm effect if that firm was neither the origin nor the destination employer of any worker transitioning in interval t . Accordingly, the resulting variability of firm effects within an interval can only be computed using the connected set of that interval. This highlights a trade-off underlying the R-AKM model: It is a very flexible model that permits firm and worker effects to vary over time but this, in turn, comes at the cost of firm effects only being identified within a relatively small set of firms.

The next subsection introduces a middle ground between the standard AKM model and the R-AKM model. This model restricts worker effects to be constant but allows firm effects to vary over time. As explained in more detail in Section 2.2.2, this extension permits the researcher to identify a firm-year effect combining multiple years of transitions as well as information on stayers.

2.2.2 The Time-Varying AKM Model (TV-AKM)

A simple extension to the standard AKM model that allows for firm-by-year heterogeneity is given by the following equation

$$y_{gt} = \alpha_g + \psi_{j(g,t),t} + \varepsilon_{gt}. \quad (15)$$

where $\{\psi_{j,t}\}$ represents a vector of firm-by-year indicators. We refer to equation (15) as the time-varying AKM model (TV-AKM).⁴ This model can be used to understand the time-series behavior of firm wage policies by, for instance, plotting the autocorrelation function of $\{\psi_{j(g,t),t}\}$.⁵ Relatedly, the TV-AKM model can be used to assess whether researchers using AKM are underestimating the importance of firms in wage setting by forcing the corresponding firm-premia to be constant over time.

A difference between TV-AKM and AKM is the role of stayers, i.e. workers that stay with the same employer over a given time window. In AKM, stayers do not contribute to the identification

⁴Bagger, Christensen and Mortensen (2014) appear to be the first that have considered a two-way model with firm-year heterogeneity. We thank a referee for pointing this out to us.

⁵Autocovariance in the firm effects can be represented as a quadratic form formed from OLS estimates of $\psi_{j(g,t),t}$. Therefore, one can use the KSS correction to bias correct these quantities for potential limited mobility biases.

of firm effects. In contrast, in TV-AKM, stayers are used to identify $\{\psi_{j,t}\}$.

To see this, consider the change in average wage at firm j between period t and $t - 1$ computed among stayers only. According to TV-AKM, this moment condition is equal to

$$\Delta_{jt} = E[y_{gt} | \text{Worker } g \text{ is a firm } j \text{ stayer}] - E[y_{g,t-1} | \text{Worker } g \text{ is a firm } j \text{ stayer}] = \psi_{j,t} - \psi_{j,t-1}. \quad (16)$$

By contrast, AKM sets $\Delta_{jt} = 0$. Interestingly, the within-worker, within-firm source of variation underlying Δ_{jt} is often also used to estimate rent-sharing elasticities (Card, Cardoso and Kline, 2015). Relatedly, Kline et al. (2019) estimate different pass-throughs between stayers and movers when looking at patent-induced shocks. This finding provides additional motivation for comparison firm effects estimated using the TV-AKM model and those based on the standard AKM model.

That stayers contribute to identification of $\{\psi_{j,t}\}$ is important also when contrasting the TV-AKM and the R-AKM models. In R-AKM, identification of firm effects within a given time interval relies solely on inter-firm transitions made by workers in that particular interval. This restricts the set of identified firm effects within and across intervals to a significant degree. On the other hand, the TV-AKM model draws on a pooled mobility network that exploits observations who are stayers. As a consequence, we expect the TV-AKM model to have a significantly higher number of identified firm-by-year effects than the R-AKM model.

3 Data, Sample Descriptive Statistics, and Growth of Wage Inequality

The estimation sample is based on quarterly wage records from all employers covered by unemployment insurance (UI) in Washington from 2002:I through 2014:IV. Each wage record consists of the quarterly earnings and work hours of a worker in a given quarter, along with an identifier of the associated employer. (A worker with multiple employers in a quarter has multiple wage records for the quarter.) Washington is unusual in requiring employers to report quarterly work hours, and it is unique in using that information to determine eligibility for UI. Washington's wage records, like

most administrative wage records, do not include workers' demographic characteristics. Further details on the data appear in the data appendix.

3.1 Construction of the Panel

We construct a linked employer-employee panel using a procedure similar to that developed by Sorkin (2018). We first identify each worker's primary employer in a quarter as the employer from which the worker had the largest share of earnings in that quarter. We then define an employment spell as at least five consecutive quarters during which a worker had earnings from the same primary employer. For each spell, we drop the first quarter (to avoid making inferences about earnings based on partial quarters of employment) and the last two quarters (to avoid making inferences based on earnings in the quarter before and the quarter of a separation). We then annualize the remaining quarterly data on earnings and hours within each calendar year, conditional on the calendar year including at least two consecutive quarters of earnings from the same primary employer. Finally, we adjust earnings by the CPI-U (indexed to 2005) and calculate the real hourly wage rate by dividing adjusted annualized earnings from the primary employer by annualized hours worked with that employer.

Figure 1 illustrates the procedure and gives some examples, described in the figure notes. The unit of observation is the worker-year, with a focus on the primary employer in a year.

3.2 Sample Descriptive Statistics and Connectivity

Table 1 displays summary statistics that are useful in assessing the econometric strategies described in Section 2. Panel (a) displays descriptive statistics for three intervals: the first and last two-year intervals we used to estimate the R-AKM model (2002–2003 and 2013–2014), and the full 2002–2014 period. For each interval, we report descriptive statistics for the largest connected set and the leave-out connected set. The former represents the largest connected set formed by worker-employer links and is the sample usually used to fit AKM wage decompositions—see for instance, Card, Heining and Kline (2013). As KSS show, unbiased estimation of variance compo-

nents requires each firm effect (or each firm-by-year effect) to be estimable after leaving out one observation. This condition requires pruning the connected set because several connected firms in the data are associated with a single mover. The sample that ensures “leave-out connectivity” is defined as the leave-out connected set.

The largest connected set for the 2002–2003 interval includes about 3.4 million worker-year observations, with about 143,000 workers who move between employers, resulting in nearly 43,000 identified firm effects. This implies that the average number of movers is approximately 3.3 movers per firm, which suggests in turn that a large fraction of firms could be associated with only one mover. The largest connected set for the 2013–2014 interval is somewhat larger than for 2002–2003—about 4 million worker-year observations, 175,000 movers, and 52,000 firms, implying on average about 3.4 movers per firm. For the full 2002–2014 period (Column 5), the largest connected set includes more than 27 million worker-year observations, nearly 2 million movers, and 252,000 firms. The longer time horizon associated with these data implies a larger mobility network and therefore a larger ratio of movers to firms, roughly 7.8.

The leave-out connected sets for the two-year interval samples (2002–2003 and 2013–2014) retain about half of the firms from the original largest connected set (Columns 2 and 4). Despite this, the leave-out samples include about 78% of the original worker-year observations and 84% of the movers, which implies that firms associated with only one mover tend to be small. Consistent with this observation and the evidence on premia paid by large firms (Oi and Idson, 1999), average log wages are higher in the leave-out connected sets than in the original largest connected sets by 6%. Also, in the leave-out samples the variances of log wages are slightly smaller than in the original largest connected sets.

The leave-out connected set for the full 2002–2014 period retains 96% of the worker-year observations, 97% of the movers, and 76% of the employers. Compared with the two-year intervals, the full observation window substantially reduces the amount of pruning needed to ensure leave-one-out connectivity.

Table 1, Panel (b) reports descriptive statistics for the sample used to fit the TV-AKM model,

which allows firm effects to vary over time. When the underlying model incorporates firm-by-year effects, the relevant bipartite network is one where vertices are formed by worker and firm-by-year identifiers, as opposed to firm identifiers. This implies a change in the definitions and sizes of the largest connected sets and the leave-out connected sets. Ultimately, though, this change appears minor: a comparison of Columns 5 and 7 shows the two samples largely overlap, and a similar conclusion holds for workers (although somewhat less for firms) in the associated leave-out samples (Columns 6 and 8). Finally, each firm is observed on average for about 5.6 years in the largest connected set, and for about 6.5 years in the leave-out connected set, suggesting substantial attrition of firms.

3.3 Growing Wage Inequality in Washington

Figure 2, Panel (a) displays the variance of real log hourly wages in Washington from 2002 to 2014. The variance of log wages increased by roughly 19%, from 0.37 to .044. This suggests that increased inequality of wages was particularly pronounced in Washington during this period: using tax data, Song, Price, Guvenen, Bloom and von Wachter (2019) find the variance of log annual earnings increased by about 4% in the United States overall during 2001–2013.

Because the data underlying our estimates and those of Song, Price, Guvenen, Bloom and von Wachter (2019) differ, this comparison could overstate differences between Washington and the United States as a whole – see the discussion in Abowd, McKinney and Zhao (2018) regarding differences between samples based on UI wage records and nationally representative samples. However, it stands to reason that earnings inequality would increase more during this period in Washington than in the United States generally because the Washington economy includes the Seattle metropolitan area, which is one of a handful of innovation hubs where wage growth has been unusually high (Moretti, 2012). From the standpoint of understanding the contributions of worker effects, firm effects, and sorting to wage inequality, Washington’s relatively large increase in wage inequality is fortunate because it presents wider scope for these effects to be observed.

Panel (b) of Figure 2 shows the evolution of different percentiles of the log wage distribution

underlying the trends shown in Panel (a). Over the 2002–2014 period, the gap between the 90th and 10th percentiles expanded by about 15 log points. This expansion occurred in three phases. Between 2002 and 2007, the increased inequality seen in Panel (a) was driven mainly by increases in the top quartile, but from 2007 to 2010 inequality accelerated, driven by a fanning out of the percentiles. After 2010, wages continued to grow for the highest quartile, and the lowest percentiles regained some of their lost ground. This was especially true for the 5th and 10th percentiles, although only the 5th wage percentile returned to its 2002 level. The relative improvement of the lower percentiles post-2010 is consistent with the relatively modest growth of inequality from 2010 to 2014 shown in Panel (a).

4 Estimating AKM at Frequent Intervals: Explaining Wage Inequality

This section presents KSS-corrected variance decompositions of log wages using a standard AKM model estimated separately for successive two-year overlapping intervals—the “rolling” AKM model (R-AKM). This R-AKM approach reveals the extent to which the rising inequality described in the previous section and shown in Figure 2 can be attributed to worker-specific effects, firm-specific wage-setting policies, and worker-firm sorting (Section 4.1). Using two-year intervals also allows us to study the cyclical nature of the variance components (Section 4.2). In Section 4.3, we examine whether inferences about the contributors to growing wage inequality change when we fail to correct for limited mobility bias; that is, we compare estimates obtained using the plug-in approach with those reported in Section 4.1 that were obtained using the KSS leave-out correction.

4.1 Explaining Rising Wage Inequality

The top panel of Table 2 displays variance decompositions obtained by fitting the standard AKM model for log wages (equation (1)) to data from the earliest (2002–2003) and the latest (2013–2014) of the two-year intervals we observe. These decompositions are bias-corrected using the KSS approach.

In the earliest interval (2002–2003), variability of worker effects explains by far the largest share (79%) of the overall variance in log wages, whereas firm effects explain only about 9%, and worker-firm sorting explains about 7%. By 2013–2014, the variability of worker effects had increased by 17% (from 0.285 to 0.333), although they explained a similar share of the variance of log wages (78%). The variability of firm effects had decreased by 22% (from 0.032 to 0.025) and by 2013–2014 explained only 6% of the variance in log wages. In contrast, the importance of worker-firm sorting nearly doubled, and explained about 12% of overall wage variation by 2013–2014 .

The rightmost two columns of Table 2 show changes between 2002–2003 and 2012–2014 in the variance of log wages and the variance components. The variance of log wages increased by almost 19% overall (from 0.360 to 0.426), and nearly 72% of this increase is explained by a rise in the variability of worker effects. But a substantial part of the increase (41%) can be attributed to rising assortativity in worker-firm matches. Variability in the wage premiums paid by firms actually decreased during the period.

To summarize, wage inequality increased substantial in Washington between 2002–2003 and 2012–2014. When we use the R-AKM model to decompose this increase, we find it was due to increases in the variation of worker-specific factors and the degree of worker-firm sorting, rather than an increase in the variation of fixed employer wage policies (which in fact decreased). The pattern we find is broadly consistent with the one in Song, Price, Guvenen, Bloom and von Wachter (2019) who fit AKM models to log earnings for seven-year intervals using 1980–2013 US administrative tax data.

4.2 Cyclicalities of the AKM Variance Components

How did the components of wage inequality evolve year-by-year over 2002–2014, a period covering the second most severe recession in US history? To address this question we move from the static comparison of Table 2 to the dynamic representation in Figure 3. Panel (a) graphs the variance decomposition shown in equation (2) for successive two-year intervals; that is, after fit-

ting a standard AKM specification separately for each interval—the R-AKM model. All variance components are bias-corrected using the KSS methodology. Panel (b) shows the changes in these variance components relative to their initial level in 2002–2003.

The variability of worker effects changed little during 2002–2006, then increased by about 20% during the Great Recession and remained at this relatively high level from 2010 on. The variability of firm effects followed a clear counter-cyclical pattern, declining by about 20% between 2002–2003 and 2007–2008, increasing by about 40% during the Great Recession, then decreasing by 45% in the post-recession years (2010–2014). In contrast, the covariance between worker and firm effects was pro-cyclical, growing by about 50% in the years leading up to the Great Recession, decreasing during the recession, then more than doubling in the post-recession years. The variability of firm effects and the importance of worker-firm sorting, then, had opposite cyclical patterns in Washington between 2002 and 2014.

To summarize, the R-AKM approach combined with the KSS correction allows us to study the cyclical nature of the AKM variance components at an annual frequency. We find that variability in worker effects increased during the Great Recession and remained relatively high. In contrast, firm effects moved counter-cyclically, increasing during the Great Recession and falling thereafter. Worker-firm sorting was pro-cyclical, falling during the Great Recession and rising dramatically thereafter.

4.3 Contrasting KSS and Plug-in Estimates

As emphasized in Section 2.2.1, limited mobility bias can be exacerbated when the AKM model is estimated separately for two-year intervals. This occurs because fewer job-to-job transitions occur in a two-year interval than in longer intervals, and these transitions are used to estimate the firm effects. As a result, we would expect that the importance of using the KSS approach to correct for limited mobility bias would be substantial when estimating AKM models using relatively short intervals and the fewer implied transitions.

To investigate this point, the bottom panel of Table 2 displays variance components estimated

using the “plug-in” approach, which does not correct for limited mobility bias and simply computes the variances and covariances over the firm-size weighted vectors of estimated worker and firm effects. As expected, based on equation (7), the plug-in method substantially overstates the importance of both worker and firm effects, and correspondingly understates the importance of worker-firm sorting. For example, in the earliest interval (2002–2003), the variance of firm effects based on the plug-in method is 50% larger than the KSS-adjusted variance (0.049 vs. 0.032). Perhaps even more striking is that, when using the plug-in approach, the estimated covariance between worker and firm effects for 2002–2003 is negative, something Andrews et al. (2008) argue is a sign of bias due to limited worker mobility.

These results illustrate the importance of using long panels to estimate AKM models unless an appropriate correction for limited mobility bias is made. The results also show why many researchers have imposed the assumption of time-invariant firm effects: the assumption allows one to pool all available transitions observed over multiple years to estimate the model, which in turn limits the degree of the limited mobility bias. The next section studies the validity of assuming fixed firm effects by contrasting the standard AKM models with an extension that allows for firm-year fixed effects.

5 Time-Varying Firm Effects: Do Firm Effects Drift?

Although using a long panel can mitigate limited mobility bias, it requires the assumption that firm effects do not vary over time. In this section, we compare estimates obtained from fitting the standard AKM model (which assumes firm fixed effects) with estimates from the TV-AKM model (which assumes firm-year fixed effects) using the pooled 2002–2014 data.

5.1 Time-Varying Firm Heterogeneity

Table 3 shows variance decompositions for log wages after estimating the standard AKM model and the TV-AKM model (equation 15) using the leave-out connected set described in Table 1, Panel (b). The KSS-corrected estimates of the standard AKM model with fixed firm effects (the

top panel) show that firm effects explain 10.6% of the total variance of log wages. When firm effects are allowed to vary over time using the TV-AKM estimator, they explain 11.6%, an 8.5% increase. The TV-AKM model also reduces the contribution of worker-firm sorting by close to 6%, and slightly increases the contribution of worker effects.

The bottom panel of Table 3 shows variance decompositions from the same two models (AKM and TV-AKM), but based on the plug-in approach rather than KSS. The differences between the AKM and TV-AKM estimates remain modest, but now allowing firm effects to vary over time increases the importance of allowing firm effects to vary. For example, using the plug-in approach, the contribution of firm effects to wage variability is about 19% higher when firm effects are allowed to vary than when they are not (14.0% versus 11.7%), whereas the difference is only 8.5% when the KSS adjustment is made. Similarly, the contribution of worker-firm sorting to wage variability is about 13% less when firm effects are allowed to vary than when they are not (14.5% versus 16.7%), whereas the difference is 5.8% when with KSS adjustment.

Evidently, in the Washington data, allowing firm effects to vary over time does not dramatically change our conclusions about the contributions of worker effects, firm effects, and worker-firm sorting to wage inequality. The share of wage variation attributable to firm effects (i.e., firm-by-year effects as opposed to fixed firm effects) increases modestly, and the share of variation attributable to worker-firm sorting decreases somewhat. And these modest differences are smaller with KSS-corrected estimates than with estimates obtained using the plug-in approach. In short, allowing firm effects to vary does not greatly change our conclusions about the sources of wage variation, and using the KSS correction reduces the importance of allowing firm effects to drift.

The analysis in Table 3 also shows that the KSS correction leads to relatively modest changes in the variance decompositions produced by the standard AKM model—compare the top and bottom panels of the AKM columns. That is, when pooling multiple years of data, mobility in the Washington data is substantial and sufficient to make limited mobility bias close to a second-order concern. In Section 7 we return to this point and show how the importance of the KSS correction varies among different time-intervals (and therefore different numbers of movers per firm or per

estimated firm-year effect). Section 8.1 will examine the degree of misspecification induced by assuming that firm effects are constant.

5.2 Autocorrelation of Firm Effects

We now examine the time-series properties of the firm-year effects estimated by the TV-AKM model. We extend the KSS-framework to bias-correct all the autocovariance and autocorrelation functions of the associated firm effects. These parameters are estimated on a balanced subsample of firms for which we can identify firm effects in each year from 2002 to 2014. Each autocovariance is then weighted by average firm employment across 2002–2014.⁶

Table 4 reports the autocovariance and autocorrelation matrices of firm effects for log wages. Firm effects appear to drift somewhat over time but they remain highly correlated even 13 years apart. Panel (a) of Figure 4 provides a visualization of the predictive power of a firm effect in 2002 on the firm effect in 2014 by showing the corresponding binned scatter plot. The KSS-adjusted slope is about 0.85 with a correlation of 0.74, and the firm effects in 2014 appear tightly clustered around the regression fit.⁷

Note that the values of the diagonal entries in the Panel (a) of Table 4 increase over time, indicating the presence of a non-stationary component in $\psi_{j,t}$. Specifically, during 2002–2014, the variance of firm effects increased from 0.0446 to 0.0582 (more than 30%). This suggests intertemporal variation in firm wage effects, and in particular, it suggests that variation in wage effects tends to increase over the life of a firm (Babina, Ma, Moser, Ouimet and Zarutskie, 2019).

Panel (b) of Figure 4, shows the autocorrelation plot of firm wage effects. We overlay the autocorrelation function of the AR(1) process that best fits the empirical autocorrelations. The resulting autoregressive parameter is 0.9765.⁸

⁶Engbom and Moser (2020) show rank weighting and unweighted autocorrelations using balanced and unbalanced samples. The results look broadly similar across these different configurations.

⁷Engbom and Moser (2020) obtain qualitatively similar results after fitting a TV-AKM model to earnings data from Sweden: the rank correlation in firm-year effects ten years apart in a balanced sample is about 0.66.

⁸Given an AR(1) process $y_t = \delta + \nu y_{t-1} + u_t$, the autocorrelation between y_t and y_{t-k} is $\rho(k) = \nu^k$. The estimate of the autoregressive parameter ν can be obtained by fitting the model $\ln(\rho(k)) = \beta \cdot k + \varepsilon(k)$, where k is the lag order, $\rho(k)$ is the k th-order autocorrelation, and ε is the regression error term. The estimate of β can be interpreted as an estimate of $\ln(\nu)$. To fit this model, the analysis uses autocorrelations reported in Table 4, Panel (b) and interpret the

Hence, the empirical autocorrelations suggest a persistent AR(1) process for firm wage effects. Figure 4 shows that the empirical autocorrelations are somewhat higher than would be predicted by an AR(1) process at lower lags (in particular, lag 1) but closely track the predicted values at higher lags.

6 Variance of Firm Effects across Models and Intervals

In this section, we further study why the variance of firm effects is lower in the R-AKM than in the TV-AKM model. This difference can be seen by comparing Tables 2 and 3. The variance of firm effects estimated using the R-AKM model with data from the last interval (2013–2014) is 0.0248 (see the top panel of Table 2), whereas the corresponding variance estimated using the TV-AKM model (and again with data from the last interval) is 0.0248 (see the top panel of Table 3). What accounts for this difference? We discuss two possible factors.

First, Tables 2 and 3 are based on different models that use different sources of variation to identify the firm (or firm-year) effects. In R-AKM, the ψ_{jt} associated with interval 2013–2014 can be identified only if firm j is part of the leave-out largest connected set observed for 2013–2014. Conversely, the TV-AKM permits pooling of stayers and movers across multiple years to identify ψ_{jt} . This implies that the set of firms used to compute the variance of firm effects reported in Table 2 is smaller (and potentially has a very different composition) than the set of firms used to compute the variance reported in Table 3.

Second, the difference between Tables 2 and 3 could be attributed to average firm premiums in 2013–2014 being different from average firm premiums in 2002–2003, which would affect the TV-AMM estimates. This difference could be due to a change in the composition of firms over the two intervals.⁹

We address the first issue—R-AKM and TV-AKM relying on different sets of firms—as fol-

antilog of $\hat{\beta}$ as an estimate of v .

⁹(Card, Heining and Kline, 2013) provide suggestive evidence that new firms in Germany are pay quite different wage premiums than existing firms, and this plays an important role in explaining the evolution of wage inequality in Germany. See also Schmieder (2021). Alternatively, the difference could be driven by “drift” of wage premiums; that is, firms changing their pay policies from 2002 to 2014.

lows. We use the estimates from the TV-AKM model estimated over 2002–2014 and report the variance of firm effects for intervals 2002–2003 and 2013–2014. We then compare the variance of the TV-AKM firm effects for those two intervals with the variance of the R-AKM firm effects for the same intervals (as reported in Table 2).

We address the second issue—potentially different average firm premiums in 2013–2014 and in 2002–2003—in a similar way. We again use the estimates from the 2002–2014 TV-AKM model and report the mean firm effects for 2002–2003 and 2013–2014. Then, to examine whether the change in the average firm effect is driven by a change in the composition of firms, we restrict the analysis to firms present both in 2002–2003 and 2013–2014.

6.1 Differences between the R-AKM and TV-AKM Firm Sets

Table 5 shows the extent to which the R-AKM and TV-AKM models rely on different sources of variation. After fitting the TV-AKM model to all available years, we identify firm-year effects for about 61,000 firms in 2002–2003 (Interval 1 in the table) and about 66,000 firms in 2013–2014 (Interval 2 in the table). In contrast, the R-AKM model identifies about 21,000 firm effects in 2002–2003 and about 26,000 firm effects in 2013–2014 (see Table 1). The identifying sets we obtain after fitting the R-AKM model separately for the two intervals are therefore less than two-fifths the size of the sets we obtain from fitting the TV-AKM model and restricting to years 2002–2003 and years 2013–2014.

Furthermore, estimates resulting from these sets appear to be different as well. The variance of the 2013–2014 firm-year effects from the TV-AKM model is about 0.0522 (Table 5), which is considerably larger than the variance of the 2013–2014 firm effects from the R-AKM model, which is about 0.0248 (top panel of Table 2). Moreover, Table 5 shows that between 2002–2003 and 2013–2014, the variance of firm effects increased, whereas in Table 2 the variance of firm effects decreased.

In summary, the R-AKM model identifies a relatively small set of firm effects, and the identified firm effects are less variable than those identified in the TV-AKM model. This highlights the key

tradeoff of the R-AKM model: The cost of allowing worker and firm effects to vary over the business cycle is the inability to identify as many firm-year effects (and the effects identified are potentially for a selected set of firms).

6.2 The Role of Firm Entry and Exit

The previous section showed that the TV-AKM and R-AKM models identify firm effects using different sets of firms, resulting in less variability of firm effects in the R-AKM than in the TV-AKM model. However, it is still possible that this difference could be driven by a changed firm premiums between 2002–2003 and 2013–2014.

To examine changes in firm effects between 2002–2003 and 2013–2014, we again use estimates from the TV-AKM model.¹⁰ Table 5 shows that average firm-year premiums increased between 2002–2003 and 2013–2014 by about 15 log points (–0.1061 to 0.0475). Using the law of total variance, we find that the difference in average firm premiums between the two intervals explains about 11 percent of the variability of firm effects reported in the first column of Table 5.¹¹

Is the difference between average firm premiums observed in 2002–2003 and 2013–2014 driven by firm entry and exit (i.e., changing composition of firms) or by changing firm-year effects of given firms? To eliminate the first possibility, the bottom panel of Table 5 restricts the results to a set of firms that are present in 2002–2003 as well as 2013–2014. The increase in average firm premiums between the two intervals in this strongly balanced set of firms is very similar to that observed for the broader set of firms: The average firm effect increases by 16 log points, from –0.1060 to 0.0518, with the difference in average firm premiums explaining around 14 percent of the pooled variance. Therefore, we conclude that the increase in average firm effects is not largely

¹⁰Footnote: The TV-AKM model allows us to meaningfully compare average firm effects across different intervals. The R-AKM model does not allow this without further assumptions because the R-AKM firm effects are expressed in relation to an interval-specific reference category. This problem is conceptually similar to the challenge of identifying the difference in firm premiums between men and women within the same firm; see the discussion in Card, Cardoso and Kline (2015).

¹¹We obtain this figure in the following way. First, for 2002–2003 and 2013–2014, we calculate the squared deviation of average firm-year effects from the simple mean of the two average firm-year effects. Second, we weight each squared deviation by the relative number of worker-year observations in each interval and sum these two numbers. Finally, we divide this between-interval variance by the pooled (2002–2003 and 2013–2014) bias-corrected variance of firm-year effects reported in the table.

driven by changes in the composition of firms over time.

In summary, the analysis suggests that average firm premiums changed between 2002–2014 largely due to changes in pay policies occurring within existing firms. However, these changes in average firm-year effects appear to explain a small share (≈ 11 – 14%) of the overall variability of firm-year effects. Most of the difference in the variability of firm effects between the R-AKM and TV-AKM models appears to result from the two models using different sets of firms to identify firm effects.

7 When is the KSS Correction Important?

A striking difference between Tables 2 and 3 is the extent to which the KSS correction changes conclusions about the relative importance of AKM variance components, depending on the length of a panel. The KSS correction is of first-order importance when using two years of data (with, on average, six movers per firm), as in Table 2. The magnitude of the KSS correction is considerably less when using 13 years of data (with, on average, 11 movers per firm). Because longer panels will, on average, have more movers per firm than shorter panels, an applied researcher may wonder whether it is the length of the panel or the number of movers per firm that is the more useful statistic in assessing the extent of limited mobility bias. In this section, we argue that the number of movers per identified firm effect is the more useful measure. Furthermore, we show how many movers per firm are required in our dataset for the plug-in and KSS-corrected estimates of variance components to align.

Figure 5 shows two kinds of relationships. The first is the relationship between the absolute value of the degree of KSS correction for a variance component (either the variance of firm effects or the correlation of worker and firm effects, on the left vertical axis) and the length of the panel used to estimate that variance component (on the horizontal axis). Variance components are estimated using a series of AKM models where the length of the panel is successively increased by one year. The first interval covers 2002–2003, the second covers years 2002–2004, the third covers

years 2002–2005, and so on.¹²

The second kind of relationship shown in Figure 5 is that between the average number of movers per number of identified firm effects (on the right vertical axis) and the length of the panel. The pattern is clear: with a panel-length of only two years, the KSS correction is crucial, especially for the correlation of worker and firm effects. As the average number of movers per firm effect increases, the degree of the KSS correction decays exponentially and stabilizes at about ten movers per firm effect.¹³

At 11 movers per firm effect (which occurs at the largest available panel length), the difference between the plug-in and KSS-corrected estimates of firm effects is about 10 percent. (In the next section, we discuss whether estimating a standard time-invariant AKM model over the full panel leads to other problems, such as model misspecification.)

Two remarks are in order. First, the number of movers per firm effect is a more useful statistic than the length of the panel in deciding whether limited mobility bias is likely to be a problem. For example, in a country like Italy, where mobility rates are lower than in Washington State (see Table 1 of KSS), it may be necessary to correct for limited mobility bias even after pooling 13 years of data.

Second, the number of movers per firm effect should be viewed as a heuristic to address the importance of the KSS bias correction, rather than a rule of thumb. The magnitude of the bias, as shown in Section 2, for the variance of firm effects is given by

$$Bias = \sum_{j=1}^J s_j V[\hat{\psi}_j] \quad (17)$$

where s_j is the size of firm j and $V[\hat{\psi}_j]$ represents the squared standard error on the OLS estimate $\hat{\psi}_j$. Clearly, if there are very few movers to estimate many of the $\hat{\psi}$ s, we expect large standard

¹²The degree of the KSS correction is defined as the absolute value of the percentage difference between the plug-in and bias-corrected estimates. That is, for the variance of firm effects, $\frac{|\hat{\sigma}_{\psi}^2 - \tilde{\sigma}_{\psi}^2|}{\tilde{\sigma}_{\psi}^2} \times 100$ where $\tilde{\sigma}_{\psi}^2$ is the uncorrected variance of firm effects and $\hat{\sigma}_{\psi}^2$ is the KSS-adjusted variance of firm effects.

¹³Bonhomme et al. (2020) also study bias-corrections for different panel lengths and obtain qualitatively similar results to these.

errors, $V[\hat{\psi}_j]$, for these firms and ultimately a significant KSS bias-correction. But $V[\hat{\psi}_j]$ can be large for other reasons (e.g., a highly heteroskedastic error term). Moreover, each (squared) standard error is ultimately weighted by firm size. If large firms have only a handful of movers, limited mobility bias will be exacerbated. In short, it is always advisable to compare plug-in and KSS-corrected estimates of the variance components when estimating AKM or similar high-dimensional linear models.¹⁴

8 Three Further Questions

In this section, we study several extensions of the main analysis. First, we show that when time-invariant firm effects are included as a continuous explanatory variable in a regression, they perform in a remarkably similar manner to time-varying firm effects. Second, we show that the autocorrelation of firm effects of new firms is much lower than that of incumbent firms. Finally, we show that, when measured 11 years apart, the autocorrelation of worker effects is similar to the autocorrelation of firm effects.

8.1 Does Allowing for Time-Varying Firm Heterogeneity Matter?

In Section 5 we showed that the variance decompositions resulting from the standard AKM model and the TV-AKM model are quite similar, and also that firm effects are highly autocorrelated. Nevertheless, by restricting firm effects to be fixed over time, the standard AKM model may still be misspecified. A formal specification test is challenging due to the high-dimensionality of the AKM and TV-AKM models. Still, it may be possible to provide suggestive evidence to inform applied researchers about the degree of misspecification stemming from restricting firm effects to be fixed over time.

We offer two types of such evidence. The first is the projection of firm effects estimated by the TV-AKM model onto firm effects estimated by the AKM model. Assuming the AKM model

¹⁴A software package that executes the KSS correction in large samples is available at <https://github.com/rsaggio87/LeaveOutTwoWay>.

represents the true data generating process, we would expect the slope of this projection to be unity, and the associated scatter plot to be highly concentrated around the 45-degree line. Panel (a) of Figure 6 shows such a projection, where the observations are worker-year weighted averages of the KSS-corrected TV-AKM and AKM firm effects within each centile bin of the AKM firm effects. The KSS-adjusted slope is 1.02, and the centile bins are tightly concentrated around the associated 45-degree line, consistent with the AKM model being correctly specified. Engbom and Moser (2020) obtain a very similar result using Swedish data.

Second, several researchers have examined the relationships between firm effects and other outcomes. For example, Bana et al. (2018) use California administrative records from 2000 to 2014 to estimate an AKM model for log earnings. They then use the estimated firm fixed effects for log earnings to gauge the importance of firm-specific earnings premiums in predicting the propensity of individuals to use public paid-leave benefits. Goldschmidt and Schmieder (2017) take a similar approach in estimating the relationship between firm effects and the wage penalty associated with outsourcing.

We take a similar approach and estimate the relationship between separation rates and firm effects, as in Card, Heining and Kline (2012). The question is whether this estimated relationship is similar when firm effects are estimated conventionally (i.e., by the standard AKM model) and by TV-AKM. In particular, we consider a worker-year weighted OLS regression of the following type

$$\text{Separation}_{it} = a + b\hat{\psi}_{j(i,t),t} + e \tag{18}$$

where Separation_{it} equals one if the worker i is employed in year t with primary employer j and is employed with a different primary employer in year $t + 1$; $\hat{\psi}_{j(i,t),t}$ is a firm effect estimated using the TV-AKM model; and e is a regression error term.

Figure 6, Panel (b) shows the results of estimating equation (18) where we compute the worker-year weighted average of the separation rates, TV-AKM firm effects, and AKM firm effects within each centile bin of the AKM firm effects. The figure suggests a minimal degree of misspecification

in the AKM model, as the overall relationships between separation rates and firm effects are similar for both fixed and time-varying firm effects. The sign of the estimate of b is negative, consistent with the evidence in Card, Heining and Kline (2012) for Germany and in Bassier, Dube and Naidu (Forthcoming) for the state of Oregon.

8.2 How Stable are the Wage Policies of New Firms?

Recent evidence suggests that pay policies of newly established firms might be quite different from those of existing firms (Schmieder, 2021). This leads to the question whether the time-series properties of firm-year effects are also different for new firms—that is, are firm premiums less persistent in new firms as those firms learn and adapt to market conditions? In this section, we analyze time-series properties of firm effects of new firms using the TV-AKM model.

Because we do not know the age of firms in our 2002–2014 panel, we define new firms as those who first appear in 2003. We then track the evolution of their firm-year effects. Figure 7 shows that the firm effect observed in the first year of the firm (2003) predicts the firm effect observed in 2014, but this predictive power is considerably less striking than was seen for all firms; see Figure 4, Panel (a). The correlation between firm effects for 2003 and 2014 is about 0.46 for new firms, whereas in the pooled sample the corresponding correlation is about 0.79 (Table 4B).

Table A1 (see the appendix) displays the autocorrelation matrix of firm effects for new firms. The correlation between the firm effect estimated in 2003 (the first year the firm is observed) and in 2004 is 0.93. This 1-year-lag autocorrelation appears to be quite variable at the beginning of the firm’s life; however, firm pay policies appear to stabilize 5–6 years after a firm’s birth. The matrix also suggests notable cyclical influences on the pay policies of new firms: during the Great Recession (2008–2009), the 1-year-lag autocorrelation drops to 0.77.

In short, firm-pay policies of new firms exhibit substantially more drift than do those of existing firms, although they appear to settle down after 5–6 years.

8.3 Do Worker Effects Drift?

The focus of this paper has been to understand the time-series properties of the firm effects. But the R-AKM model allows both firm and worker effects to change between intervals, so it is natural to ask whether worker effects drift. In this section, we briefly investigate the stability of worker effects and discuss potential directions for future research.

Figure 8 displays a binned scatter plot of worker effects in 2013–2014 against worker effects in 2002–2003. (The estimated worker effects underlying the plot come from the models used to estimate the decompositions shown in Table 2.) Worker effects for 2002–2003 strongly predict worker effects estimated more than 10 years later, with a correlation of about 0.70, similar to what we found for firm effects. The lower end of the bin scatter is flat, suggesting that workers with low fixed effects in 2002–2003 experienced a common increase by 2013–2014 that was independent of their initial fixed effect.

To summarize, worker and firm fixed effects appear to show a similar degree of persistence. Worker effects in 2002–2003 strongly predict worker effects observed in 2013–2014, although this prediction is not perfect—worker effects do exhibit some drift over time. This could be the result of multiple factors, such as accumulation of human capital ((Bagger et al., 2014), employer-learning (Altonji and Pierret, 2001), or changes to an individual’s outside option (Caldwell and Harmon, 2019). Applying the R-AKM approach might be fruitful in explaining changes in the gender gap over the age profile (Card, Cardoso and Kline, 2015, Di Addario et al., 2020), although this is a question that we cannot address with these data. In general, extending the AKM model to allow for worker trends as well as worker fixed effects would appear to be a useful avenue for future research.

9 Conclusions

We have addressed two main questions. First, do firm effects “drift” over time? Our analysis suggests that by and large, firm effects are quite stable and that little is lost in assuming firm effects

to be constant over time. These conclusions are based on three pieces of evidence from Washington administrative wage records: First, when estimated over 13 years, fixed and time-varying firm effects explain similar shares of variance decompositions. Moreover, we observe almost perfect correlation between fixed and time-varying firm effects. Second, the autocorrelation function of firm effects appears to be generated by a process resembling a highly persistent AR(1) (specifically, with an autoregressive parameter equal to about 0.98). The implication is that, even 13 years apart, firm effects remain highly autocorrelated. Third, when using estimated firm effects as explanatory variables, as in several recent papers, fixed and time-varying firm effects generate nearly identical predictions.

Finding that firm effects are relatively stable provides a justification for pooling multiple years of data to estimate AKM firm effects. The observed stability of firm effects also has broader implications: The time-invariance of firm pay policies suggests that firm effects for pay reflect primarily permanent differences in compensation policies. Whether these differences arise from inter-firm differences in value-added per worker and associated rent sharing, from differences in worker preferences for the nonwage characteristics of employers, or other reasons remains an open question.

The second main finding is that correcting for limited mobility bias is important when estimating the AKM model over short time intervals. Specifically, without the KSS bias correction, AKM variance decompositions calculated over two-year intervals attribute too much importance to worker and firm effects and too little to worker-firm sorting. Conversely, AKM variance decompositions calculated over a 13 years long panel are relatively similar with and without the KSS bias correction. This similarity holds whether or not firm effects are fixed or allowed to vary over time.

Our results suggest at least two avenues for future work. One is to reconcile different economic models that give rise to firm wage premia with the stochastic process of the time-varying firm effects reported here. For example, one could extend the static model in Card et al. (2018) to allow for richer dynamics. A second could be to examine the generality of the observed pattern of firm effects over the business cycle in Washington State, and to better understand the mechanism behind

this pattern within existing or new macroeconomic models.

References

- Abowd, John M., Francis Kramarz and David N. Margolis. 1999. “High Wage Workers and High Wage Firms.” *Econometrica* 67(2):251–333. (document)
- Abowd, John M., Francis Kramarz, Paul Lengeremann and Sébastien Pérez-Duarte. 2004. Are Good Workers Employed by Good Firms? A Test of a Simple Assortative Mating Model for France and the United States. Unpublished manuscript. 1, 2
- Abowd, John M., Kevin L. McKinney and Nellie L. Zhao. 2018. “Earnings Inequality and Mobility Trends in the United States: Nationally Representative Estimates from Longitudinally Linked Employer-Employee Data.” *Journal of Labor Economics* 36(S1):S183–S300. 3.3
- Altonji, Joseph G and Charles R Pierret. 2001. “Employer learning and statistical discrimination.” *The quarterly journal of economics* 116(1):313–350. 8.3
- Andrews, Martyn J., Len Gill, Thorsten Schank and Richard 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(3):673–697. 1, 2, 4.3
- Babina, Tania, Wenting Ma, Christian Moser, Paige P. Ouimet and Rebecca Zarutskie. 2019. Pay, Employment, and Dynamics of Young Firms. Working Paper 21 Federal Reserve Bank of Minneapolis, Opportunity and Inclusive Growth Institute. 5.2
- Bagger, Jesper, Bent Jesper Christensen and Dale T. Mortensen. 2014. Productivity and wage dispersion: Heterogeneity or Misallocation. Unpublished manuscript. 4
- Bagger, Jesper, Fran Fontaine, Fabien Postel-Vinay and Jean-Marc Robin. 2014. “Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics.” *American Economic Review* 104(6):1551–96. 8.3

- Bana, Sarah, Kelly Bedard, Maya Rossin-Slater and Jenna Stearns. 2018. Unequal use of social insurance benefits: The role of employers. Technical report National Bureau of Economic Research. 2, 8.1
- Bassier, Ihsaan, Arindrajit Dube and Suresh Naidu. Forthcoming. “Monopsony in Movers: The Elasticity of Labor Supply to Firm Wage Policies.” *Journal of Human Resources* . 8.1
- Benedetto, Gary, John Haltiwanger, Julia Lane and Kevin McKinney. 2007. “Using Worker Flows to Measure Firm Dynamics.” *Journal of Business & Economic Statistics* 25:299–313. 15
- Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad and Bradley Setzler. 2020. How Much Should we Trust Estimates of Firm Effects and Worker Sorting? Technical report National Bureau of Economic Research. 13
- Caldwell, Sydnee and Nikolaj Harmon. 2019. “Outside options, bargaining, and wages: Evidence from coworker networks.” *Unpublished manuscript, Univ. Copenhagen* . 8.3
- Card, David, Ana Rute Cardoso, Jörg Heining and Patrick Kline. 2018. “Firms and Labor Market Inequality: Evidence and Some Theory.” *Journal of Labor Economics* 36(S1):S13–S70. 1, 9
- Card, David, Ana Rute Cardoso and Patrick Kline. 2015. “Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women.” *Quarterly Journal of Economics* 131(2):633–686. 1, 2.2.2, 10, 8.3
- Card, David, Francesco Devicienti and Agata Maida. 2014. “Rent-sharing, Holdup, and Wages: Evidence from Matched Panel Data.” *Review of Economic Studies* 81(1):84–111. 1
- Card, David, Jörg Heining and Patrick Kline. 2012. Workplace Heterogeneity and the Rise of West German Wage Inequality. Working Paper 18522 National Bureau of Economic Research. 8.1, 8.1
- Card, David, Jörg Heining and Patrick Kline. 2013. “Workplace Heterogeneity and the Rise of

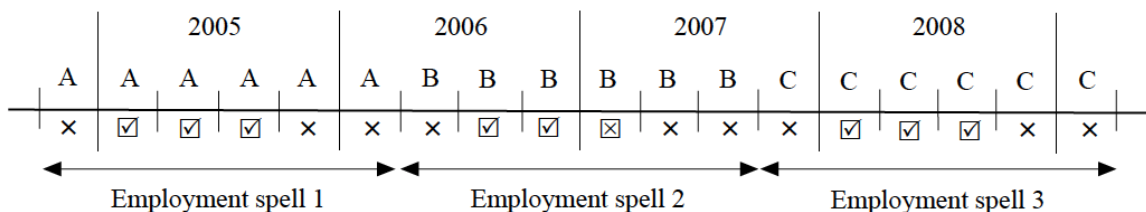
- West German Wage Inequality.” *Quarterly Journal of Economics* 128(3):967–1015. 1, 2, 2.2.1, 3.2, 9, A.2, A.3
- Cattaneo, Matias D, Michael Jansson and Whitney K Newey. 2018. “Inference in linear regression models with many covariates and heteroscedasticity.” *Journal of the American Statistical Association* 113(523):1350–1361. 2.1.2
- Di Addario, Sabrina, Patrick Kline, Raffaele Saggio and Mikkel Solvsten. 2020. “It ain’t where you’re from, it’s where you’re at: hiring origins, firm heterogeneity, and wages.”. 8.3
- Engbom, Niklas and Christian Moser. 2020. “Firm pay dynamics.” *Available at SSRN 3531250* . 1, 6, 7, 8.1
- Garin, Andrew, Filipe Silvério et al. 2019. How Responsive are Wages to Demand within the Firm? Evidence from Idiosyncratic Export Demand Shocks. Working paper. 2
- Goldschmidt, Deborah and Johannes F. Schmieder. 2017. “The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure.” *Quarterly Journal of Economics* 132(3):1165–1217. 1, 2, 8.1
- Guiso, Luigi, Luigi Pistaferri and Fabiano Schivardi. 2005. “Insurance within the Firm.” *Journal of Political Economy* 113(5):1054–1087. 1
- Kline, Patrick, Neviana Petkova, Heidi Williams and Owen Zidar. 2019. “Who Profits from Patents? Rent-sharing at Innovative Firms.” *Quarterly Journal of Economics* 134(3):1343–1404. 1, 2, 2.2.2
- Kline, Patrick, Raffaele Saggio and Mikkel Sølvssten. 2020. “Leave-out estimation of variance components.” *Econometrica* 88(5):1859–1898. (document), 1, 2, ??
- Krueger, Alan B. and Lawrence H. Summers. 1988. “Efficiency Wages and the Inter-industry Wage Structure.” *Econometrica* 56(2):259–293. 1, 2

- Lachowska, Marta, Alexandre Mas and Stephen A. Woodbury. 2018. How Reliable Are Administrative Reports of Paid Work Hours? Unpublished manuscript. A.1
- Lachowska, Marta, Alexandre Mas and Stephen A Woodbury. 2020. “Sources of displaced workers’ long-term earnings losses.” *American Economic Review* 110(10):3231–66. 1, 2, A.1, A.2, A.3
- Lamadon, Thibaut, Magne Mogstad and Bradley Setzler. 2019. Imperfect Competition, Compensating Differentials and Rent Sharing in the US Labor Market. Working Paper 25954 National Bureau of Economic Research. 3
- MacKinnon, James G and Halbert White. 1985. “Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties.” *Journal of econometrics* 29(3):305–325. 2.1.2
- Moretti, Enrico. 2012. *The New Geography of Jobs*. Houghton Mifflin Harcourt. 3.3
- Oi, Walter Y. and Todd L. Idson. 1999. “Firm Size and Wages.” *Handbook of Labor Economics* 3:2165–2214. 3.2
- Schmieder, Johannes. 2021. “Establishment Age and Wages.” *Mimeo* . 9, 8.2
- Schmieder, Johannes F., Till von Wachter and Joerg Heining. 2018. The Costs of Job Displacement Over the Business Cycle and Its Sources: Evidence from Germany. Working paper Boston University: Mimeo. 1
- Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom and Till von Wachter. 2019. “Firming Up Inequality.” *Quarterly Journal of Economics* 134(1):1–50. 1, 2, 3.3, 4.1
- Sorkin, Isaac. 2018. “Ranking Firms using Revealed Preference.” *Quarterly Journal of Economics* 133(3):1331–1393. 1, 3.1, A.2
- Van Reenen, John. 1996. “The Creation and Capture of Rents: Wages and Innovation in a Panel of UK Companies.” *Quarterly Journal of Economics* 111(1):195–226. 1

White, Halbert. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica* pp. 817–838. 2.1.2

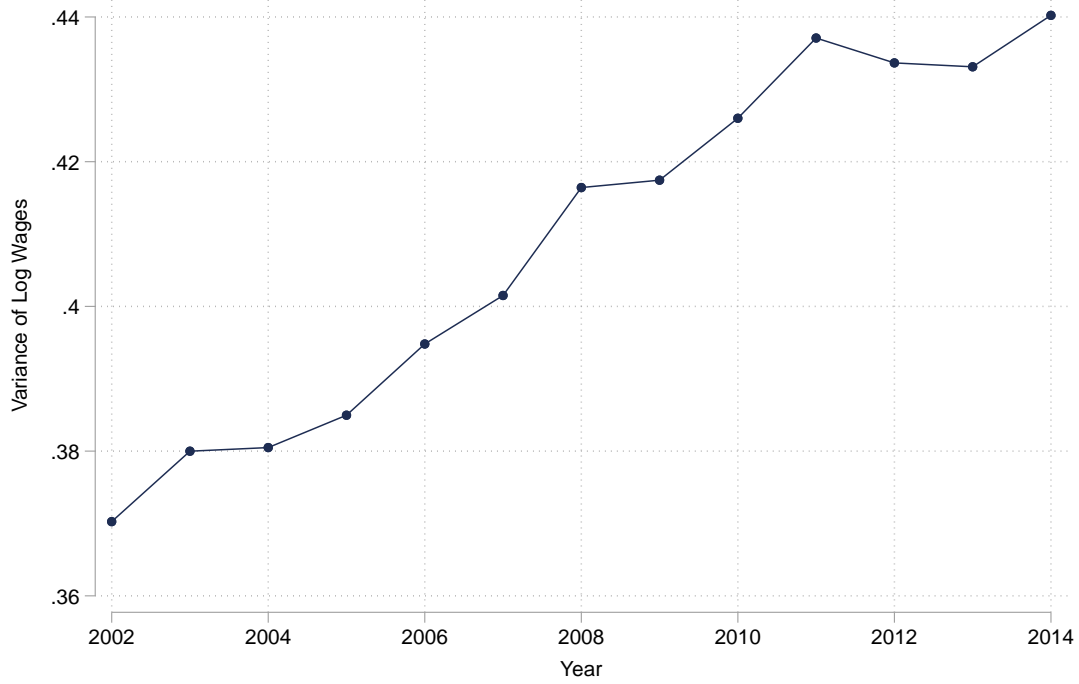
10 Figures and Tables

Figure 1: Construction of the Analysis Sample

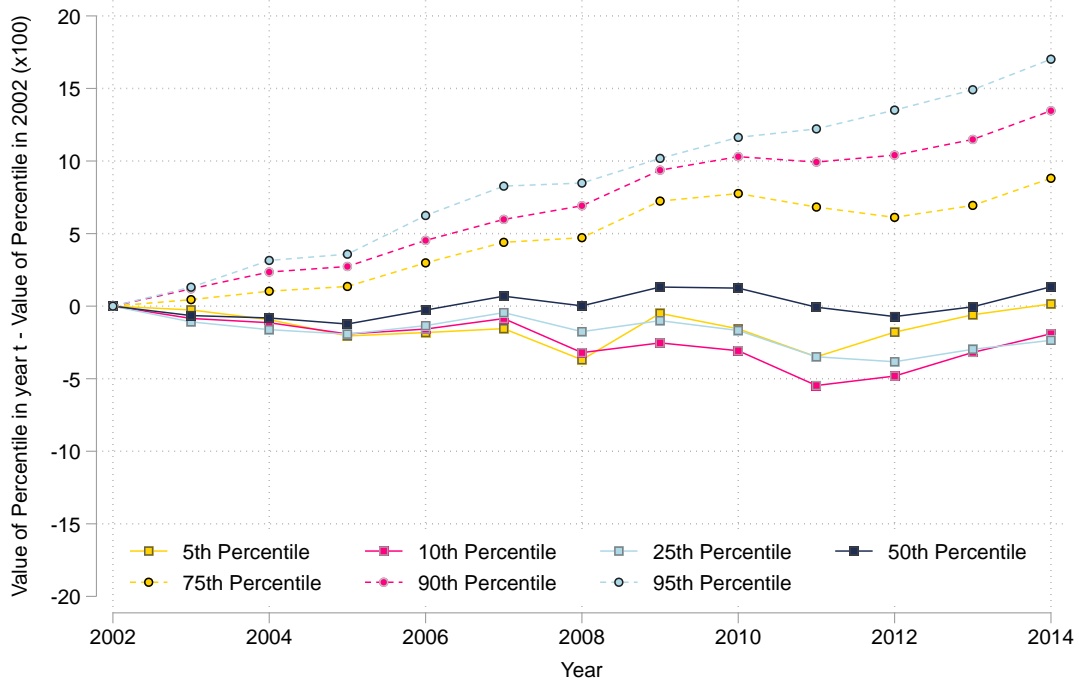


Notes: The figure shows three hypothetical employment spells with three different employers (A, B, and C), each of which has the minimum five quarters required to be included in the analysis sample. The first quarter and last two quarters of each employment spell (denoted by ×) are dropped from the analysis, and outcomes from the remaining quarters are then annualized for each calendar year, conditional on the calendar year including at least two consecutive quarters of earnings from the same primary employer. For example, outcomes for 2005 (Employment spell 1) and 2008 (Employment spell 3) are obtained by averaging the outcomes for the first, second, and third quarters of 2005 (or 2008) and multiplying by four. (The quarters used in the calculations are denoted by ☑.) Outcomes for 2007 (part of Employment spell 2) are excluded because 2007 does not include two consecutive quarters that can be used under the selection criteria (that is, after excluding the first quarter and last two quarters of each employment spell). As a result, the data from 2007:Q1 (denoted by ☒) are not used.

Figure 2: Trends in Wage and Earnings Inequality



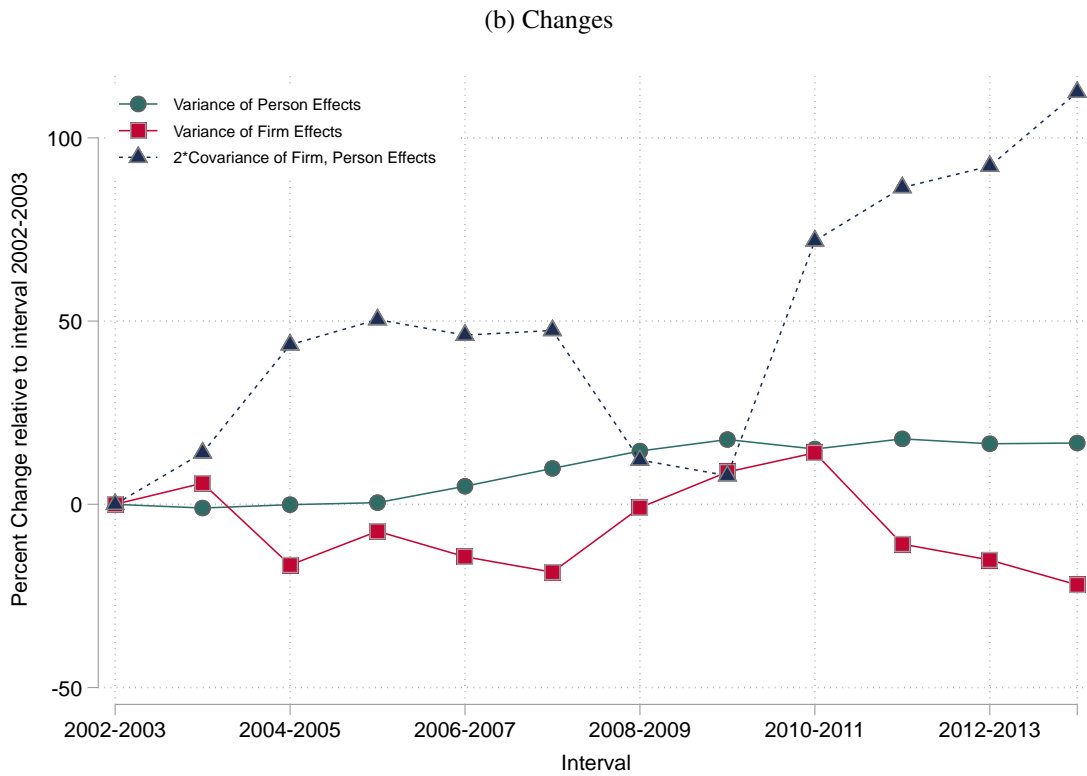
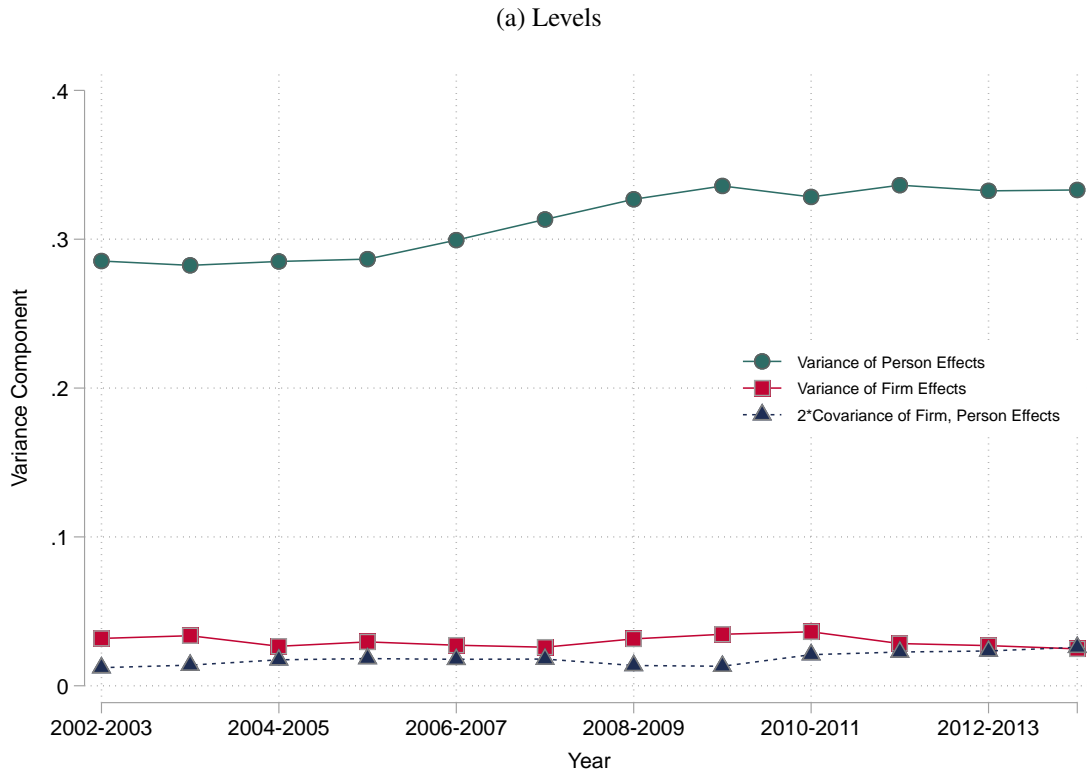
(a) Variance of Log Wages



(b) Trends in Percentiles

Note: Panel (a) plots the unconditional variance of log hourly wages in the Washington administrative data. Panel (b) plots the evolution of the log hourly wages percentiles, where each percentile has been deviated from the value of the same percentile in 2002.

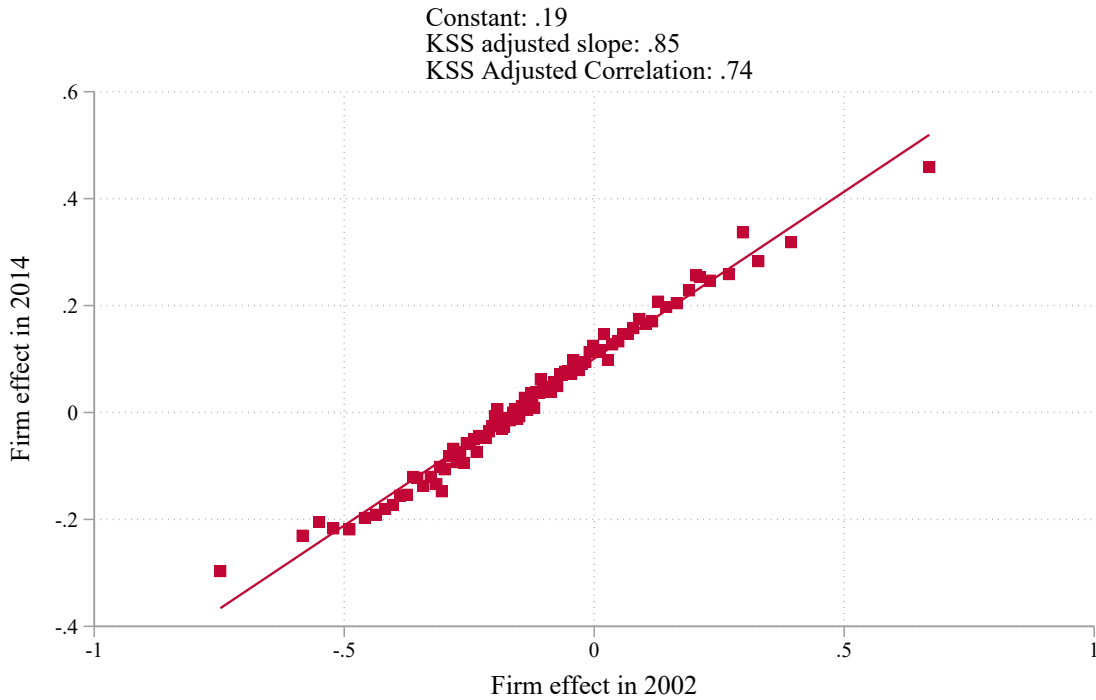
Figure 3: Evolution of AKM Variance Components



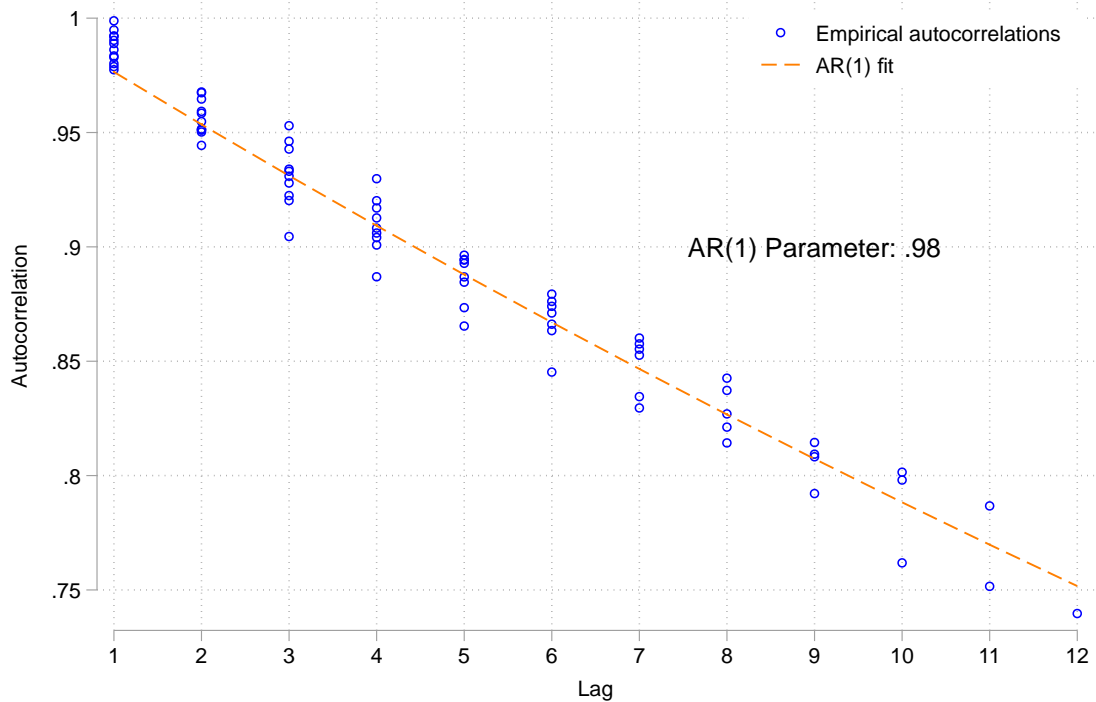
Note: Panel (a) displays the variance components of the AKM model after fitting an AKM separately in each two-year adjacent interval. Each variance component is corrected using the bias correction of Kline, Saggio and Sølsten (2020) based on leaving a worker-firm match out. Panel (b) displays the changes in these variance components relative to their initial level from 2002–2003.

Figure 4: Autocorrelation of Firm Effects

(a) Scatter Plot

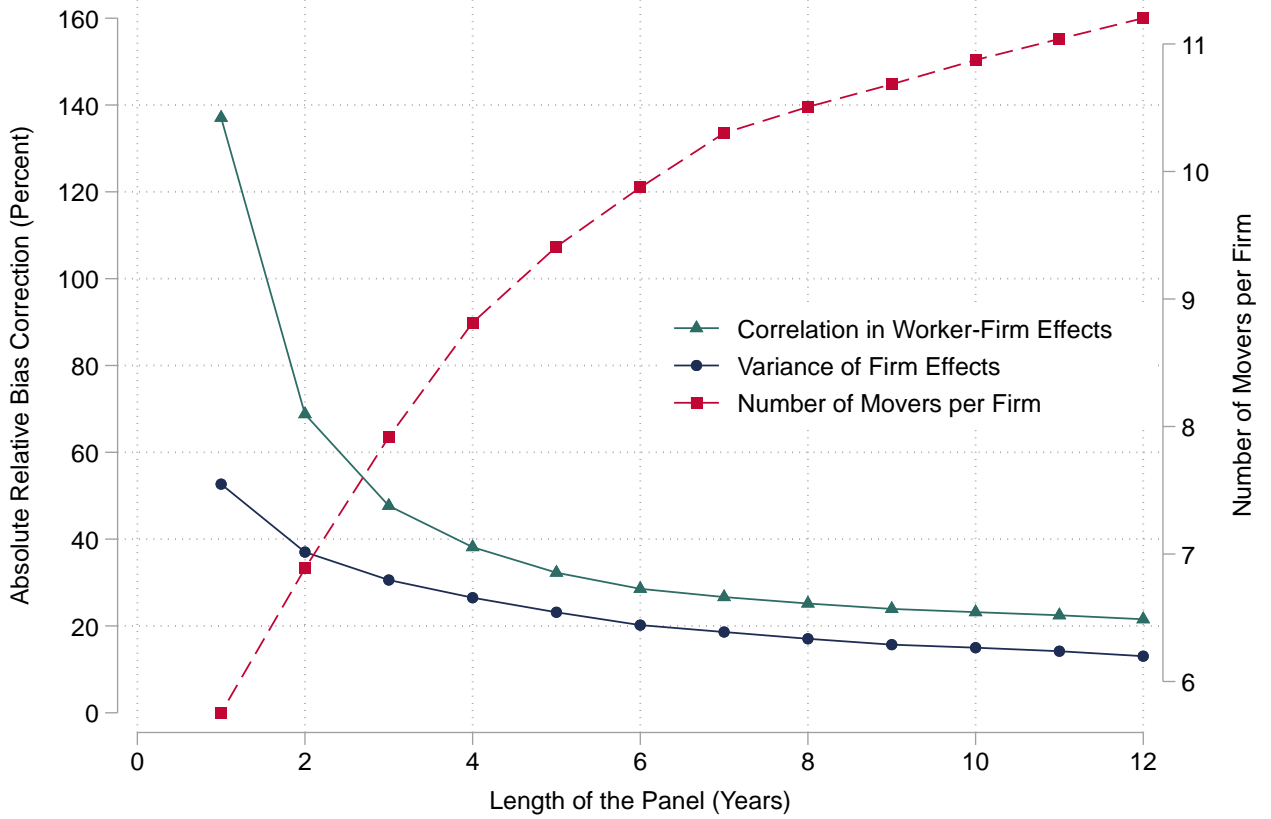


(b) Autocorrelation



Note: Panel (a) displays a binned scatter plot of the relationship between the 2014-firm effect and the 2002-firm effect, where both set of effects are estimated using the TV-AKM model. Each square represents a centile of the firm effect estimated in 2002. In each bin we then report the average firm effect in 2014 and 2002. Both the regression slope as well as associated correlation are KSS-adjusted. Panel (b) plots the estimated autocorrelations in Table 4B (blue dots) and the autocorrelation function of the AR(1) process $\psi_{j,t} = 0.98\psi_{j,t-1} + \zeta_{jt}$.

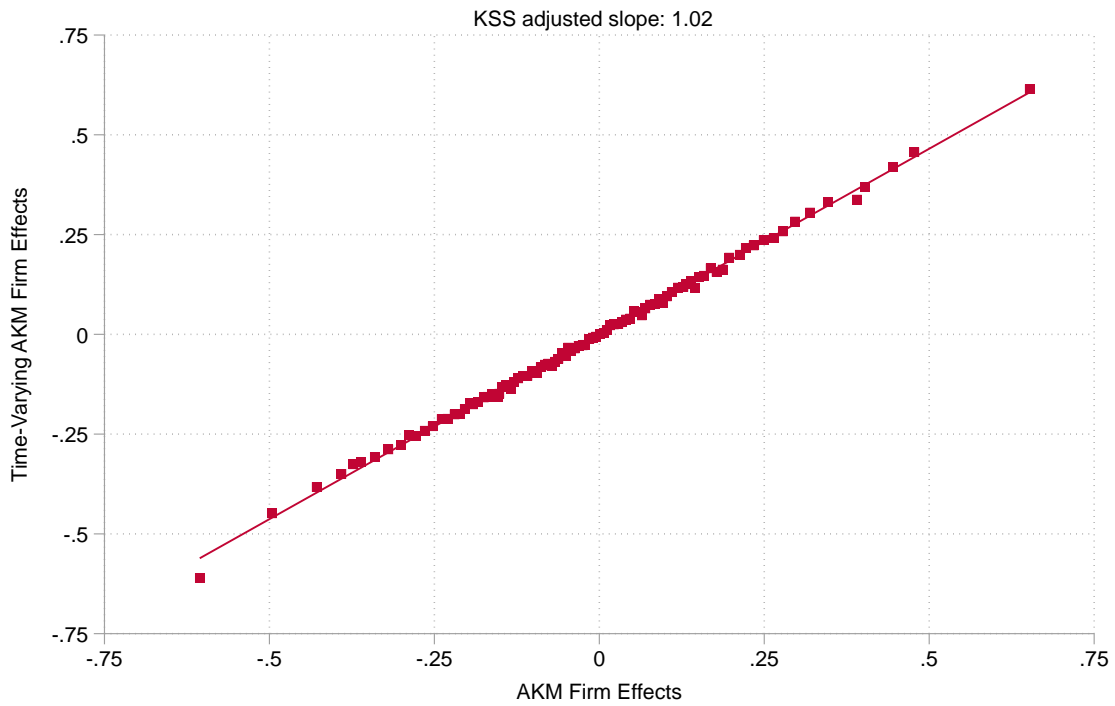
Figure 5: KSS Correction Over Different Panel Lengths



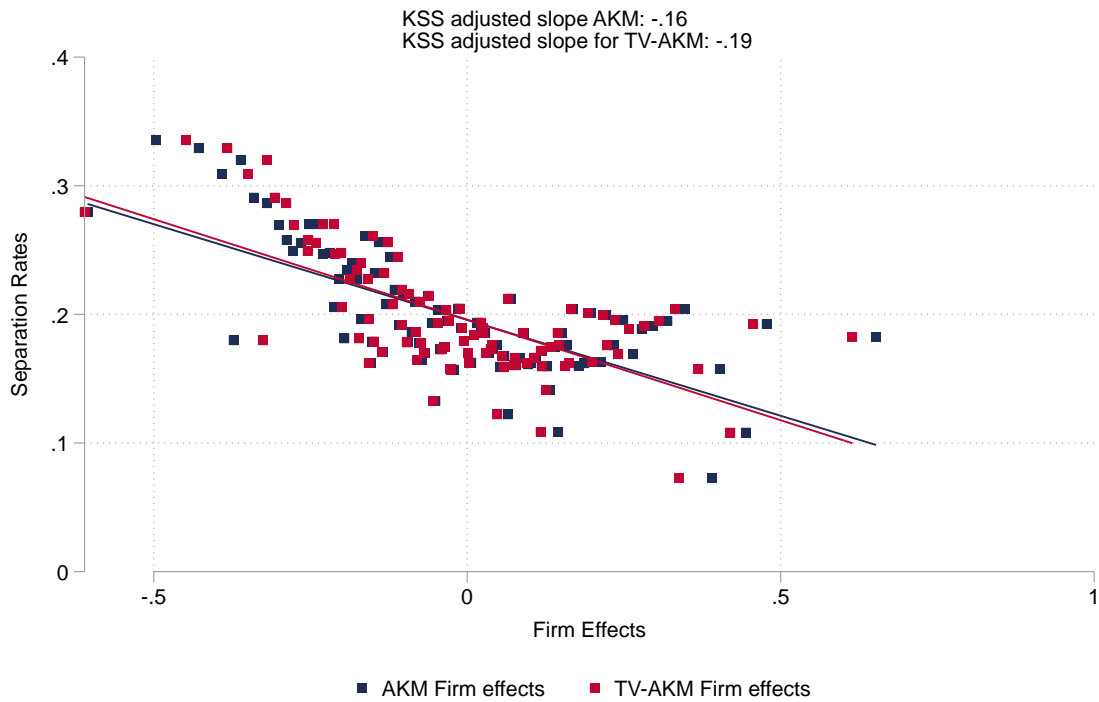
Note: The figure plots the degree of the KSS bias correction of the variance of firm effects and the correlation of worker and firm effects based from an AKM model fitted over panels of varying lengths (2002–2003, 2002–2004, etc). The right axis reports the associated number of movers per firm effect for a given panel. The degree of the KSS correction is reported as the percentage difference between the KSS and PI estimate. For example, for the variance of firm effects, the degree of bias correction is given by $\frac{|\hat{\sigma}_{\psi}^2 - \tilde{\sigma}_{\psi}^2|}{\tilde{\sigma}_{\psi}^2} \times 100$, where $\tilde{\sigma}_{\psi}^2$ is the uncorrected variance of firm effects and $\hat{\sigma}_{\psi}^2$ is the KSS adjusted variance of firm effects. KSS correction is estimated leaving a worker-firm match out.

Figure 6: Does Allowing for Time-Varying Firm Heterogeneity Matter in Practice?

(a) TV-AKM vs. AKM

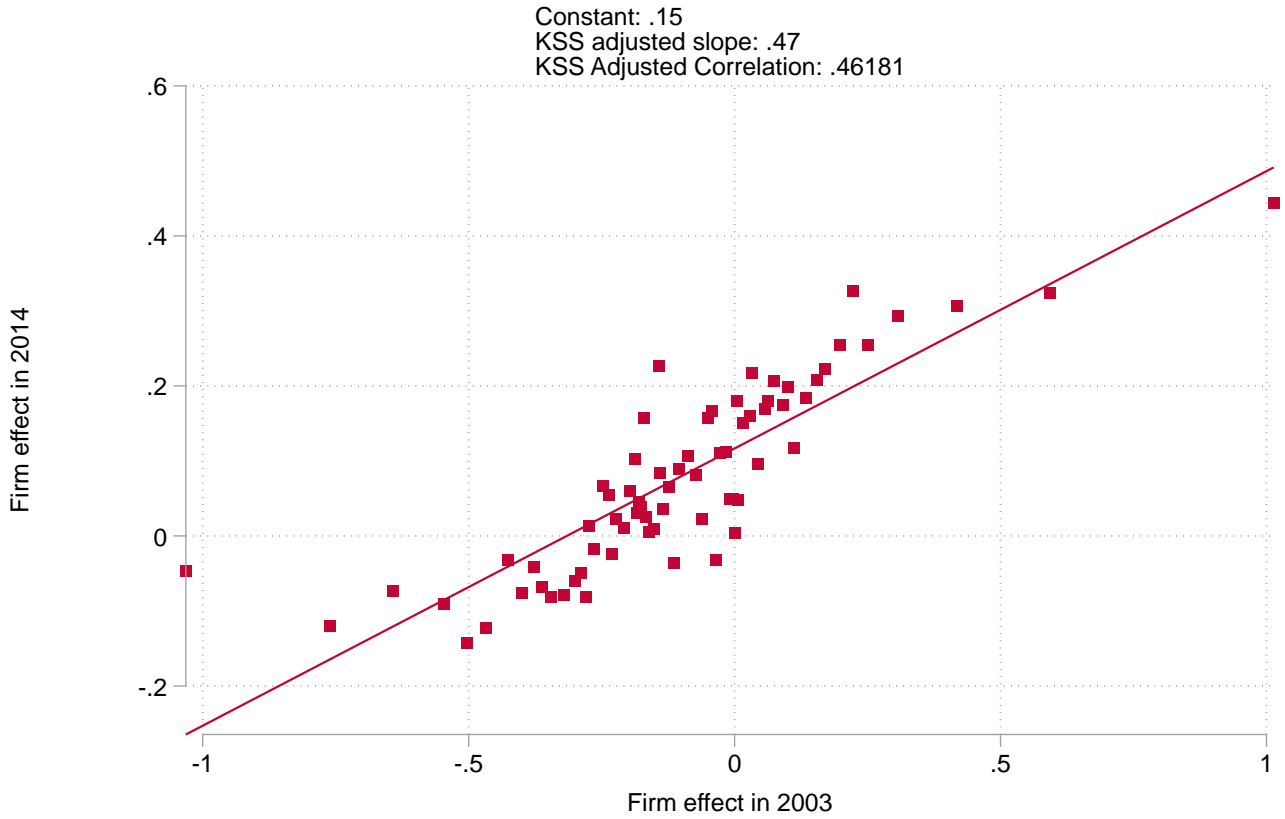


(b) KSS



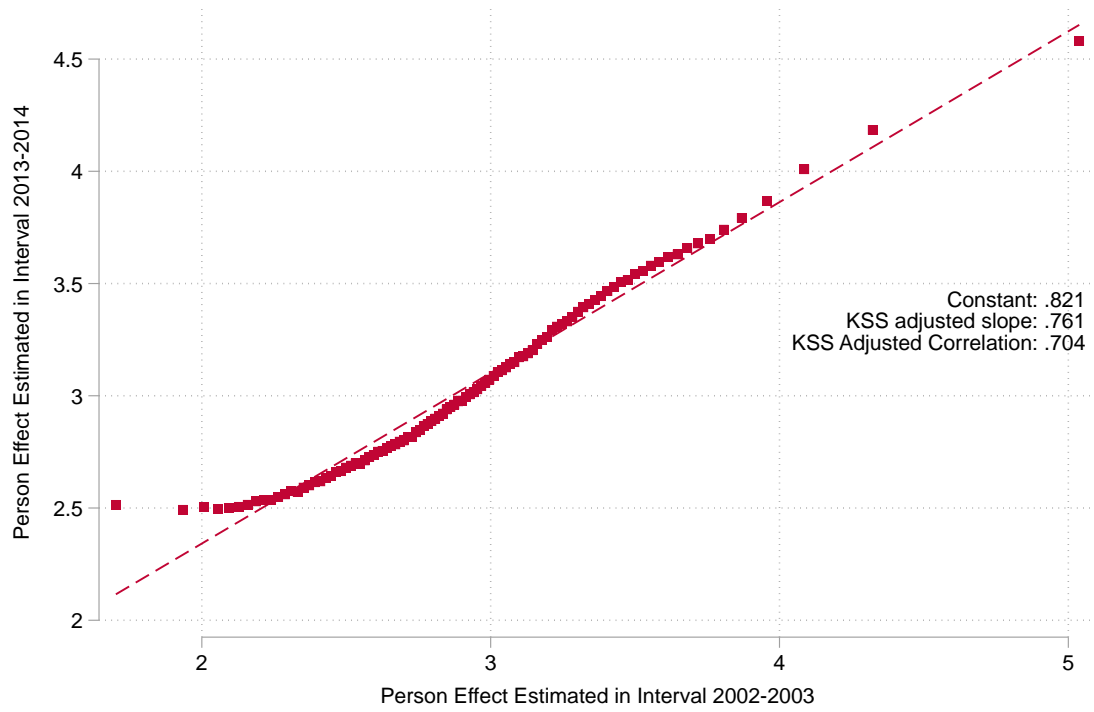
Note: Panel (a) reports the worker-year averages of the AKM firm effects and the TV-AKM firm effects within each centile bin of the AKM effects, see the text for further details. Both set of effects are estimated in a pooled dataset which stacks together all available years from 2002 to 2014. In Panel (b) we report worker-year averages of separation rates, TV-AKM firm effects and

Figure 7: Autocorrelation of Firm Effects for New Firms



Note: This figure displays a binscatter plot of the firm-year effects for “new” firms defined as employers that appear for the first time in the data in 2003. Specifically, we contrast the firm-year effect in the first year of the life of the firm – 2003 – against its corresponding value in the last available year – 2014. Each square represents a centile of the firm effect estimated in 2003. In each bin we then report the average firm effect in 2014 and 2003. Both the regression slope as well as associated correlation are KSS-adjusted.

Figure 8: Do Worker Effects Drift?



Note: The figure displays a binned scatter plot of the relationship between the worker effects estimated using AKM on data from 2002–2003 and worker effects estimated using AKM on data from 2013–2014. Each square represents a centile of the worker effect estimated in 2002–2003. In each bin we then report the average worker effect in 2013–2014 and 2002–2003. The slope of the underlying regression has been corrected using the KSS methodology. Similarly for the reported correlation.

Table 1: Summary Statistics

	<i>Panel (a): AKM Model</i>				<i>Panel (b): TV-AKM Model</i>			
	<i>Interval: 2002-2003</i>		<i>Interval: 2013-2014</i>		<i>Interval: 2002-2014</i>		<i>Interval: 2002-2014</i>	
	Largest Connected Set [1]	Leave Out Connected Set [2]	Largest Connected Set [3]	Leave Out Connected Set [4]	Largest Connected Set [5]	Leave Out Connected Set [6]	Largest Connected Set [7]	Leave Out Connected Set [8]
# of Person Year Obs	3,437,465	2,667,634	4,075,427	3,224,620	27,662,224	26,233,816	27,608,217	24,743,671
# of Movers	143,485	119,984	175,578	148,363	1,964,604	1,884,040	1,962,768	1,718,563
# of Firms	42,844	20,849	51,888	26,055	252,571	168,186	250,788	118,751
# of Firms by year							1,402,407	768,558
Mean Log Wages	3.03	3.09	3.06	3.12	3.02	3.03	3.02	3.04
Variance of Log Wages	0.38	0.36	0.44	0.43	0.41	0.41	0.41	0.41

Note: This table reports the descriptive statistics on the different samples used in the analysis. Panel (a) summarizes the samples used when fitting standard AKM models where firm effects are not allowed to vary over time. Panel (b) describes samples used to fit an AKM model where firm effects are allowed to vary over time. The largest connected set in Panel (a) is defined as the largest sample where all the firms are connected by workers' moves. The largest connected set in Panel (b) is defined as the largest sample where all the firm-by-year combinations are connected by workers' moves. The leave-out connected set is the largest sample where all firms (or all firms-by-year combinations, for Panel-b) remain connected to each other after leaving out a worker from the sample. A mover is defined in both Panel (a) and Panel (b) as an individual that changed firm within a given time frame. All statistics on wages are person-year weighted. See text for details.

Table 2: Variance Decomposition based on Rolling AKM

	<i>Interval 2002-2003</i>		<i>Interval 2013-2014</i>		<i>Change from 2002-2003 to 2013-2014</i>	
	Variance Component	Share of Total (%)	Variance Component	Share of Total (%)	Variance Component	Share of Total (%)
Variance of Log Wages	0.3598	100.00%	0.4264	100.00%	0.0666	100.00%
<i>Variance Decomposition: KSS</i>						
Variance of Person Effects	0.2854	79.31%	0.3331	78.12%	0.0478	71.69%
Variance of Firm Effects	0.0318	8.84%	0.0248	5.82%	-0.0070	-10.50%
2*Cov of Person, Firm Effect	0.0243	6.75%	0.0516	12.11%	0.0273	41.05%
<i>Variance Decomposition: Plug-In</i>						
Variance of Person Effects	0.3108	86.40%	0.3554	83.34%	0.0445	66.82%
Variance of Firm Effects	0.0488	13.57%	0.0393	9.22%	-0.0095	-14.28%
2*Cov of Person, Firm Effect	-0.0085	-2.36%	0.0239	5.61%	0.0324	48.62%

Note: All variance decomposition parameters are based on an AKM model fitted separately within each time interval and are calculated in the corresponding leave one out connected set described in Table 1, Panel (a). Plug-in reports the variance components without adjusting for sampling error in the estimated person and firm effects. KSS adjusts each variance component using the leave out approach detailed by Kline, Saggio and Sølvesten (2020). The last two columns report the change in a corresponding row over time. That change is then scaled by the change in the variance of log wages reported in first row. All variance components are person-year weighted. The R-AKM controls for year fixed effects. Source: WA administrative records.

Table 3: Variance Decomposition - Pooled Data 2002-2014

	<u>AKM</u>		<u>TV-AKM</u>	
	Variance Component	Share of Total (%)	Variance Component	Share of Total (%)
Variance of Log Wages	0.4069	100.00%	0.4069	100.00%
<u>Variance Decomposition: KSS</u>				
Variance of Person Effects	0.2495	61.31%	0.2533	62.24%
Variance of Firm Effects	0.0433	10.63%	0.0470	11.55%
2*Cov of Person, Firm Effects	0.0745	18.31%	0.0703	17.27%
<u>Variance Decomposition: Plug-In</u>				
Variance of Person Effects	0.2568	63.10%	0.2594	63.75%
Variance of Firm Effects	0.0478	11.74%	0.0569	13.99%
2*Cov of Person, Firm Effects	0.0678	16.66%	0.0590	14.50%

Note: All variance decomposition parameters are calculated in the corresponding leave one out connected set described in Table 1, Panel (b) and are person-year weighted. TV-AKM corresponds to an AKM model where firm effects are allowed to vary over-time. In the AKM model, firm effects are instead constant over-time. Both models allow for time fixed effects. Plug-in reports the variance components without adjusting for sampling error in the estimated person and firm effects. KSS adjusts each variance component using the bias correction procedure of Kline, Saggio and Sølvssten (2020) by leaving a worker-firm match out. Source: WA administrative records.

Table 4A: Autocovariance of Firm Effects - Log Wages

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
2002	0.0446	0.0428	0.0400	0.0378	0.0381	0.0369	0.0371	0.0362	0.0364	0.0368	0.0366	0.0363	0.0378
2003		0.0427	0.0412	0.0389	0.0390	0.0378	0.0381	0.0373	0.0375	0.0381	0.0380	0.0378	0.0394
2004			0.0402	0.0393	0.0391	0.0378	0.0380	0.0370	0.0373	0.0379	0.0377	0.0374	0.0389
2005				0.0391	0.0401	0.0387	0.0389	0.0376	0.0374	0.0379	0.0375	0.0372	0.0387
2006					0.0413	0.0407	0.0407	0.0392	0.0394	0.0401	0.0399	0.0397	0.0414
2007						0.0408	0.0416	0.0397	0.0394	0.0402	0.0401	0.0401	0.0418
2008							0.0431	0.0422	0.0418	0.0425	0.0425	0.0421	0.0440
2009								0.0426	0.0427	0.0432	0.0432	0.0429	0.0447
2010									0.0448	0.0459	0.0459	0.0457	0.0476
2011										0.0484	0.0491	0.0487	0.0507
2012											0.0516	0.0509	0.0527
2013												0.0524	0.0548
2014													0.0586

Table 4B: Autocorrelation of Firm Effects - Log Wages

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
2002	1.00	0.98	0.94	0.90	0.89	0.87	0.85	0.83	0.81	0.79	0.76	0.75	0.74
2003		1.00	0.99	0.95	0.93	0.91	0.89	0.87	0.86	0.84	0.81	0.80	0.79
2004			1.00	0.99	0.96	0.93	0.91	0.89	0.88	0.86	0.83	0.81	0.80
2005				1.00	1.00	0.97	0.95	0.92	0.89	0.87	0.83	0.82	0.81
2006					1.00	0.99	0.96	0.93	0.92	0.90	0.86	0.85	0.84
2007						1.00	0.99	0.95	0.92	0.90	0.87	0.87	0.86
2008							1.00	0.98	0.95	0.93	0.90	0.88	0.88
2009								1.00	0.98	0.95	0.92	0.91	0.89
2010									1.00	0.99	0.95	0.94	0.93
2011										1.00	0.98	0.97	0.95
2012											1.00	0.98	0.96
2013												1.00	0.99
2014													1.00

Note: This table computes the autocovariance and autocorrelation function of the firm-year effects estimated from the TV-AKM model. All auto-covariance and autocorrelation parameters reported in the table above are computed for the sample of firms that are alive in each year from 2002-2014 and are weighted using the average number of workers associated with a given firm from 2002 to 2014. Variance components are corrected using the leave-out methodology of Kline, Saggio and Sølvssten (2020).

Table 5: Entry / Exit of Firms and Changes to Firm-Wage Premia

	Pooling Interval 1 and Interval 2	Interval 1 (2002-2003)	Interval 2 (2013-2014)
<i>All Firms</i>			
Variance of Log Hourly Wages	0.4024	0.3704	0.4301
# of Firms	97,563	60,990	65,832
# of Person-Year Observations	7,363,510	3,427,668	3,935,842
Average Firm-Year Effect	-0.0240	-0.1061	0.0475
Variance of Firm-Year Effects (Bias-Corrected)	0.0527	0.0412	0.0522
<i>Balanced Firms</i>			
Variance of Log Hourly Wages	0.3947	0.3652	0.4198
# of Firms	25,556	25,556	25,556
# of Person-Year Observations	5,100,806	2,364,054	2,736,752
Average Firm-Year Effect	-0.0213	-0.1060	0.0518
Variance of Firm-Year Effects (Bias-Corrected)	0.0450	0.0418	0.0539

Note: This table reports estimates based from a TV-AKM model that control for year fixed effects fitted over the entire panel, i.e. years 2002--2014. The first column reports statistics when pooling the interval 2002-2003 and 2013-2014 together. Second column report statistics for 2002-2003 only while the third column focuses on 2013-2014. In panel (b), the statistics of interest are reported within the subsample of balanced firms, i.e. firms that are present both in 2002-2003 as well as 2013-2014. All statistics are person-year weighted. Variance of firm-year effects are corrected using the KSS methodology leaving a worker-firm match out.

A Data Appendix

This appendix first describes the data used in the estimation of firm effects for hourly wages used in the main text. It then describes the sample restrictions imposed on the estimation sample. Finally, we highlight some considerations specific to working with single-state administrative wage records.

A.1 Further Description of the Data

The data used in this paper are based on quarterly administrative wage records maintained by the Employment Security Department of Washington State to administer the state's unemployment insurance (UI) system. The available quarterly data provide information on earnings and work hours of all workers employed by UI-covered employers in the state between 2002–2014. Workers who drop out of the labor force, become self-employed, work in the underground economy, or move out of state will not appear in the records. This is because self-employed workers are not covered by UI, underground earnings are not reported, and out-of-state earnings will be picked up in the earnings records of another state; see also the discussion below.

UI-covered employers in Washington are required to report each worker's quarterly earnings and work hours, which allows us to construct an hourly wage rate in each quarter for most workers in Washington's formal labor market. Lachowska, Mas and Woodbury (2018) examine the reliability of the Washington hours data and conclude they are of high quality. Each worker's quarterly record also includes an employer identifier and the employer's four-digit North American Industry Classification System (NAICS) code, making it possible to construct employment at both the employer and industry level. The data do not include geographic locations of the employers or the workers. For further discussion of Washington administrative wage records, see Lachowska, Mas and Woodbury (2020).

The available data do not include demographic characteristics for all of the workers—we only observe demographic characteristics (date of birth, gender, race/ethnicity, and education) for workers who separated from their employer, claimed UI, and received it. Because these workers rep-

resent only about 30 percent of the panel, restricting the analysis to these workers would break the connectivity of the set of identified firm effects. Hence, the analysis here does not control for demographic characteristics.

A.2 Sample Restrictions

As in Lachowska, Mas and Woodbury (2020), we impose several restrictions on the estimation sample, dropping the following:

- Workers with more than 9 employers in a year
- Workers with annual earnings less than \$2,850 (in 2005 dollars) and workers with calculated hourly wage rates \leq \$2.00/hour (in 2005 dollars) (following Card, Heining and Kline 2013; Sorkin 2018)
- Workers who worked fewer than 400 hours in the year

A.3 Additional Considerations

First, although we use the terms “employer” and “firm” interchangeably, they are not always the same. The employer is the entity from which UI payroll taxes are collected and is the unit of observation in the wage records. For firms with a single establishment, and for firms with multiple establishments all located in Washington, the employer is also the firm. (In some cases, a multi-establishment firm may be divided into more than one employer.) For firms with multiple establishments some of which are located outside Washington, the employer covers only the firms’ establishments located in Washington.

Second, an employer identification (ID) number may disappear if the employer becomes inactive or reorganizes in some way—through merger, acquisition, spinoff, breakout, or other reason. The available data do not include an employer “successor file,” so we cannot distinguish among these cases with certainty,¹⁵ but like Card, Heining and Kline (2013), we take the view that a new

¹⁵Using information on worker flows, Benedetto et al. (2007) develop probabilistic approach to identifying employers that have undergone an employer identification number change due to merger, acquisition, spin off, or breakout.

employer ID likely implies reorganization and new employment policies, so it makes sense to treat an employer ID change as the creation of a new entity and to estimate a separate (new) employer fixed effect. As Card, Heining, and Kline (2013) point out, treating assignment of a new ID to an existing employer leads to a loss of efficiency but no bias.

Third, we cannot follow workers who move out of Washington. To examine how such attrition might influence the estimated AKM employer fixed effects, Lachowska, Mas and Woodbury (2020) reestimate the AKM model after dropping random subsamples of 30 percent and 50 percent of employers from the full AKM sample described above. The resulting firm effects are highly correlated with those obtained in the full AKM analysis, suggesting that interstate migration is not a substantial source of bias in estimating the AKM fixed effects.

B Computational Appendix

In this appendix we provide additional details about the implementation of the KSS methodology.

B.1 Leave-Out Connected Set

Lemma 1 in KSS states that unbiased estimation of variance components requires that the leave-out OLS coefficient $\hat{\beta}_{-i}$ exists when dropping any observation i from the sample. In the AKM context, this requirement is satisfied by conducting estimation on the “leave-out connected set”. This is defined as the largest connected set where all the firms remain connected after removing a particular worker from the sample, see KSS for details.

One can easily extend the notion of connected set and leave-out connectivity for the TV-AKM model. In this model, the underlying network is formed by workers and firm-by-year identifiers, as opposed to firm-only identifiers as in the original AKM formulation. Intuitively, the TV-AKM model treats each firm-year combination as a separate firm. It follows that identification of $\{\alpha_g, \psi_{j(g,t),t}\}$ from 15 requires all the firm-year combinations to be connected.¹⁶ Leave-out con-

¹⁶As for AKM, the scale of the firm-year effects is not identified. It is necessary therefore to set one particular firm-year effect to zero within a given connected set to ensure that the design matrix of the TV-AKM model to be of full rank. This normalization does not affect the resulting variance components.

nectivity requires that each firm-year combination remains connected after removing a particular worker from the sample.

In general, one should expect leave-out connectivity when applied to a TV-AKM model to imply further pruning of the data when compared to a leave-out connected set based off an AKM model. Figure 9 illustrates this latter point with a simple example. It shows a network that satisfies the definition of leave-one-out connectedness when estimating a model with time-invariant firm heterogeneity. However, if we were to estimate the full set of firm-by-year effects in this example, the associated definition of leave-one-out connectedness would no longer be satisfied.

B.2 Leave-Out Strategy

Equation (13) show the KSS heteroskedastic-unbiased (HU) estimate of the standard errors which is used to derive the KSS correction for the variance components of interest. As discussed in KSS-Remark 4, when fitting an AKM model with only two time periods, like in the R-AKM model, the KSS estimate of the variance component is robust not only to unrestricted heteroskedasticity of the regression error term but also arbitrary serial correlation of ε_{g1} and ε_{g2} when estimating the variance of firm effects and the covariance of worker and firm effects.

When pooling multiple years of data, it may be important to allow for dependence in the error terms in $\{\varepsilon_g\}$. The KSS approach can be extended to allow for dependence of the error term in such scenarios by leaving, instead of a single observation, a cluster of observations when computing the leave-out OLS estimate β . In our analysis, we report KSS bias-corrected variance of firm effects and covariance of worker, firm effects that are robust to dependence of the error term within a worker-firm match, i.e. we run the KSS correction by leaving out a match out.

Note that when leaving out a worker-firm match, the worker effects of workers who are observed with only one firm are not identified. Therefore, our resulting KSS bias-corrected variance of worker effects is computed after leaving a single observation out. We report HU estimates for the resulting auto-correlation functions.

B.3 Computation of the Statistical Leverages

The main computational bottleneck of the KSS correction is computation of the statistical leverages of the associated linear model: $P_{ii} = x_i'(\sum_{i=1}^n x_i x_i')^{-1} x_i$. Exact computation of these quantities for a dataset the size of the WA administrative data is unfeasible.

We therefore rely on a variation of the Johnson-Lindestrauss approximation developed by KSS. This methodology is inspired by random projection algorithms popular in computer science and is used to provide a fast but highly accurate estimate of P_{ii} . KSS show that the Johnson-Lindestrauss approximation cut computation time by a factor of 100 while introducing an approximation error of roughly 10^{-4} . All the baseline bias-corrected variance decompositions from the AKM, R-AKM, and TV-AKM models are computed by relying on the available package for KSS, available here: <https://github.com/rsaggio87/LeaveOutTwoWay>.

C Additional Tables and Figures

Table A1-A: Autocovariance of Firm Effects New Firms - Log Wages

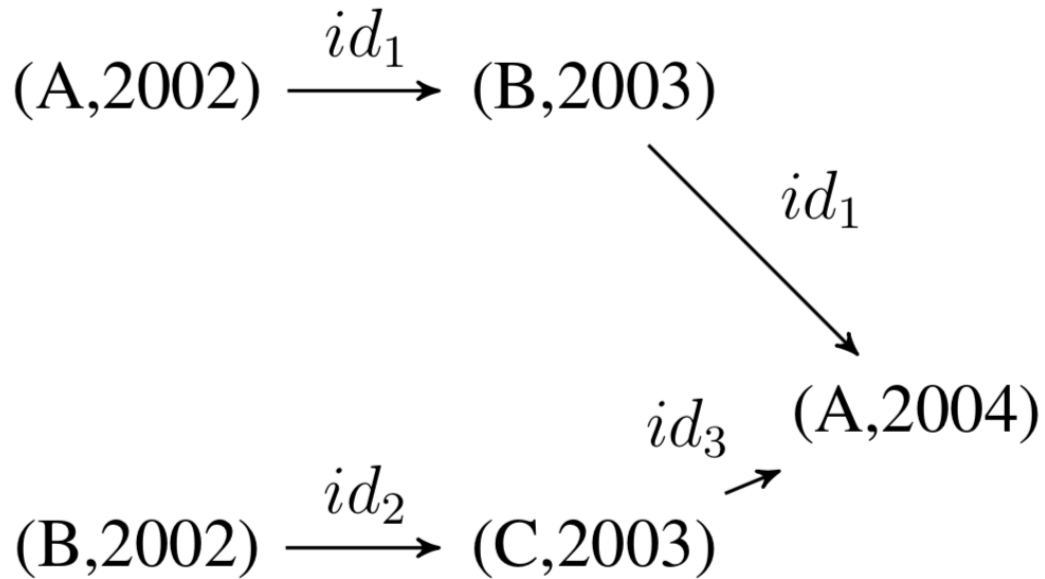
	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
2003	0.0539	0.0464	0.0402	0.0405	0.0366	0.0422	0.0314	0.0300	0.0294	0.0262	0.0257	0.0251
2004		0.0453	0.0400	0.0402	0.0367	0.0436	0.0321	0.0314	0.0317	0.0291	0.0289	0.0282
2005			0.0392	0.0410	0.0369	0.0476	0.0332	0.0323	0.0329	0.0317	0.0311	0.0312
2006				0.0447	0.0432	0.0623	0.0379	0.0367	0.0376	0.0369	0.0360	0.0363
2007					0.0426	0.0664	0.0386	0.0376	0.0392	0.0386	0.0380	0.0387
2008						0.1434	0.0587	0.0570	0.0602	0.0611	0.0604	0.0632
2009							0.0404	0.0408	0.0425	0.0420	0.0410	0.0411
2010								0.0437	0.0469	0.0459	0.0442	0.0444
2011									0.0499	0.0511	0.0492	0.0494
2012										0.0534	0.0525	0.0527
2013											0.0516	0.0533
2014												0.0547

Table A1-B: Autocorrelation of Firm Effects New Firms - Log Wages

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
2003	1.00	0.93	0.87	0.82	0.76	0.48	0.67	0.62	0.56	0.49	0.49	0.46
2004		1.00	0.95	0.89	0.83	0.54	0.75	0.70	0.67	0.59	0.60	0.57
2005			1.00	0.98	0.90	0.63	0.83	0.78	0.74	0.69	0.69	0.67
2006				1.00	0.99	0.78	0.89	0.83	0.80	0.76	0.75	0.74
2007					1.00	0.85	0.93	0.87	0.85	0.81	0.81	0.80
2008						1.00	0.77	0.72	0.71	0.70	0.70	0.71
2009							1.00	0.97	0.95	0.91	0.90	0.88
2010								1.00	1.00	0.95	0.93	0.91
2011									1.00	0.99	0.97	0.95
2012										1.00	1.00	0.98
2013											1.00	1.00
2014												1.00

Note: This table computes the autocovariance and autocorrelation function of the firm-year effects estimated from the TV-AKM model for new firms, that is, employers that appear for the first time in the data in 2003. All auto-covariance and autocorrelation parameters reported in the table above are computed for the sample of firms that are alive in each year from 2003-2014 and are weighted using the average number of workers associated with a given firm from 2003 to 2014. Variance components are corrected using the leave-out methodology of Kline, Saggio and Sølvssten (2020).

Figure 9: Leave-One-Out Connected Set with Time-Varying Firm Heterogeneity



Note: The figure depicts a connected bipartite graph where vertices are given by firm-by-year combinations—e.g. firm A in year 2003—and edges are formed whenever a given worker (id_1, id_2, id_3) transition from one vertex to another. The above graph does not satisfy the condition of leave-one-out connected set as dropping one worker disconnects the graph. If one defines vertices as a single firm instead of a firm-by-year combination, the resulting graph would satisfy the condition of a leave-one-out connected set.