Sectoral Shocks and Labor Market Dynamics: A Sufficient Statistics Approach

Jaeeun Seo* Ryungha Oh†

December 2024

Abstract

We develop a sufficient statistics approach to evaluate the impact of sectoral shocks on labor market dynamics and welfare. Within a broad class of dynamic discrete choice models that allows for arbitrary persistent worker heterogeneity, we show that knowledge of steady-state intersectoral gross worker flows over various time horizons is sufficient to evaluate labor supply responses to shocks and their welfare consequences. We also establish analytically that assuming away persistent worker heterogeneity—a common practice in the literature—leads to overestimation of steady-state worker flows, resulting in systematic biases in counterfactual predictions. As an illustration of our approach, we revisit the consequences of the rise of import competition from China. Using US panel data to measure steady-state worker flows, we conclude that labor reallocation from manufacturing to non-manufacturing is significantly slower, and the negative welfare effects on manufacturing workers are much more severe than those predicted by models without persistent worker heterogeneity.

Keywords: Sectoral Shocks, Labor Market Dynamics, Sufficient Statistics, Comparative Advantage, Worker Heterogeneity, Self-Selection, Dynamic Discrete Choice.

*Northwestern University; Email: `jaeeun.seo@northwestern.edu`
†Northwestern University; Email: `ryungha.oh@northwestern.edu`

1. Introduction

Labor markets in the United States, as well as in many other countries, have been subject to a variety of shocks, from globalization to the rise of automation, oil price shocks, and the Covid-19 pandemic. Although these shocks differ in many ways, they all have one thing in common: their effects tend to be highly asymmetric across sectors, potentially creating both winners and losers among workers. How much do winners gain and losers lose? Can workers exposed to a negative shock in one sector avoid, or at least mitigate, its adverse consequences by moving to another sector? And if so, what determines the extent of this reallocation and the time it takes?

The goal of this paper is to shed light on these questions. The premise of our analysis is that both workers' exposure to shocks and their subsequent sectoral mobility depend on their comparative advantage across sectors. If comparative advantage is weak or highly transient, we expect frequent sector changes of workers, resulting in small welfare losses or even gains for negatively exposed workers. If, instead, comparative advantage is strong and persistent, we expect many workers to remain stuck in the negatively affected sector for a long time, leading to more severe negative welfare consequences. The key idea of this paper is that whether comparative advantage is weak or strong is revealed by workers' propensity to switch sectors *prior to shocks*—i.e., steady-state gross worker flows. Thus, we can use such information as sufficient statistics to evaluate the impact of sectoral shocks on labor market dynamics and welfare. In this paper, we formalize this intuition and demonstrate its importance both theoretically and empirically.

We focus on a broad class of dynamic discrete choice models that allows for arbitrary time-invariant heterogeneity across workers. At each point in time, workers decide in which sector to work. They are subject to sector-switching costs and idiosyncratic shocks that are independently drawn each period from an extreme value distribution, as in the canonical dynamic discrete choice framework (e.g., Artuç, Chaudhuri, and McLaren, 2010). However, workers may have time-invariant differences in sector-specific productivity (or, more generally, in the utility they derive from being in a particular sector) and in sector-switching costs. We impose no restriction on these differences, in line with the general Roy model (e.g., Heckman and Honore, 1990). Within this general environment, we establish two theoretical results.

First, we develop a novel sufficient statistics approach that yields valid counterfactual predictions without estimating the sources or extent of worker heterogeneity. Aside from the discount factor and a single parameter that governs the dispersion of the idiosyncratic shocks, this approach requires only one input: steady-state gross intersectoral worker flows over different time horizons.[1] We show that this information is sufficient to construct sectoral welfare changes and dynamic labor reallocation in response to sectoral shocks, up to a first-order approximation. Steady-state worker flows summarize the effect of heterogeneity of workers on their mobility, and this is precisely what we need in order to predict the dynamic effect of sectoral shocks. We start by focusing on the labor supply side, examining the effect of exogenous changes in sectoral wages. However, we also demonstrate that when we augment the model with the labor demand side and endogenize

[1] Formally, we need steady-state probabilities that workers switch from sectors i to j after n periods, for all sectors i and j and all values of n.

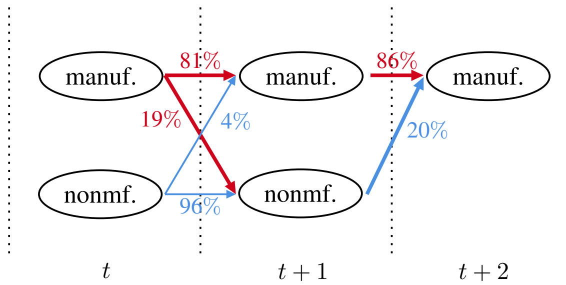


Figure 1. 1-Year and 2-Year Worker Flows in the Data

Notes: Arrows between years t and $t+1$ represent steady-state 1-year worker flows between manufacturing and non-manufacturing sectors. They represent $\Pr(s_{t+1} = \text{manuf.}|s_t = \text{manuf.}) = 1 - \Pr(s_{t+1} = \text{nonmf.}|s_t = \text{manuf.}) = 0.81$ and $\Pr(s_{t+1} = \text{nonmf.}|s_t = \text{nonmf.}) = 1 - \Pr(s_{t+1} = \text{manuf.}|s_t = \text{nonmf.}) = 0.96$. Arrows between years $t+1$ and $t+2$ provide additional information needed to compute the 2-year staying probability for the manufacturing sector. They represent $\Pr(s_{t+2} = \text{manuf.}|s_{t+1} = s_t = \text{manuf.}) = 0.86$ and $\Pr(s_{t+2} = \text{manuf.}|s_{t+1} = \text{nonmf.}, s_t = \text{manuf.}) = 0.20$. Assuming that the economy was in a steady state between years 1980 and 2000, these worker flows are computed by pooling all observations of the NLSY79 data over this period.

wage changes, the same set of sufficient statistics, combined with knowledge of the labor demand side, can be used to perform counterfactual exercises.

Second, we show analytically how persistent worker heterogeneity shapes the consequences of sectoral shocks. We begin by characterizing the systematic bias in steady-state worker flows implied by the canonical dynamic discrete choice model without persistent worker heterogeneity (hereafter, the canonical model). Ignoring heterogeneity and the resulting self-selection would lead to underestimation of the long-run probabilities of workers remaining in the same sector. Intuitively, the canonical model ignores the fact that workers who have self-selected into a sector are more likely to choose the same sector in subsequent periods. More importantly, this bias, combined with the sufficient statistics result, implies systematic biases in the counterfactual predictions of the canonical model. In particular, the canonical model always underestimates the welfare losses of adversely affected workers and overestimates the speed of labor reallocation for given exogenous changes in sectoral wages. Moreover, we show that when wages are endogenously determined, underestimation of welfare losses is likely to be compounded by overestimation of labor reallocation. These findings suggest the importance of incorporating persistent worker heterogeneity in evaluating the consequences of sectoral shocks, which is quantified in the remainder of the paper.

The next part of our paper provides empirical estimates of our sufficient statistics for the United States. We first use the panel information in the National Longitudinal Survey of Youth 1979 (NLSY79) dataset to compute worker flows over various time horizons. In line with our theoretical finding, we find that the steady-state worker flows observed in the data are inconsistent with those predicted by the canonical model without persistent worker heterogeneity. For example, the canonical model underestimates the probabilities of workers choosing the same sector after ten years by more than a factor of two. Simply controlling for workers' demographic and socioeconomic characteristics—such as gender, education, race, and age—explains little of this underestimation. This suggests that the persistent heterogeneity that drives this underestimation is mostly within these characteristics and unobserved to the econometrician.

To illustrate the inconsistency, we plot the 1-year and 2-year worker flows observed in the NLSY data in Figure 1. A detailed description of the data and calculations will be provided in Section 4, along with additional analysis. The figure shows, for example, that 81% of typical manufacturing workers remain in manufacturing after 1 year. The canonical model, which assumes away persistent heterogeneity, necessarily implies that workers who choose to stay in the manufacturing sector are as likely to stay in the following year as the typical manufacturing worker. However, the data reveal that these workers exhibit a higher probability of staying again in the following year (86% > 81%). Similarly, workers who have self-selected into manufacturing in year t are more likely to choose it again in year $t + 2$ even when they choose non-manufacturing sectors in year $t + 1$ (20% > 4%). As a result of these discrepancies, the canonical model underestimates the probability of workers choosing the manufacturing sector again after 2 years.

We then turn to estimation of the dispersion of the idiosyncratic shock. This estimation is based on the observation that the response of sectoral employment to wage shocks depends solely on the dispersion parameter and the sufficient statistics, independent of the specific details of worker heterogeneity. Thus, this parameter can be estimated by measuring the response of sectoral employment, conditional on our sufficient statistics. We put this idea into practice by extending the standard Euler equation approach in the literature to allow for arbitrary worker heterogeneity.

In the final part of our paper, we combine our empirical estimates with the sufficient statistics result to revisit two applications in the literature. First, we apply our findings to a hypothetical trade liberalization exercise of Artuç, Chaudhuri, and McLaren (2010), in which the economy experiences an unexpected permanent drop in manufacturing prices. This stylized exercise clearly illustrates how the failure to match worker flows across sectors at different horizons can lead to biases in counterfactual predictions. Second, as a more realistic application, we revisit an extensively studied topic: the impact of the rise in China's import competition on US labor markets. Following Caliendo, Dvorkin, and Parro (2019), we introduce a richer labor demand side that features international and intranational trade, input-output linkages, and multiple production inputs. The results demonstrate that labor reallocation following the China shock is significantly smaller, by up to 20%, and the negative welfare effects on manufacturing workers are more severe than those predicted by the canonical model. In the absence of persistent worker heterogeneity, workers initially employed in the manufacturing sector in all states appear to benefit, on average, from the China shock because the model predicts that they can easily move to positively affected sectors. However, this is no longer the case when we do account for persistent worker heterogeneity. The welfare gains of manufacturing workers are close to zero even when they are positive, and manufacturing workers in seven states experience welfare losses.

Related Literature. This paper is related to several strands of the literature. A large body of empirical literature studies the labor market impact of shocks that exhibit asymmetric effects across sectors, such as globalization (Goldberg and Pavcnik, 2007); automation (Acemoglu and Restrepo, 2020); the Covid-19 pandemic (Chetty et al., 2020); and oil price shocks (Keane and Prasad, 1996). Of particular relevance to our application is the literature that examines the impact of the rise in China import competition on US labor markets (e.g., Autor, Dorn, and Hanson, 2013; Autor et al., 2014; Acemoglu et al., 2016; Pierce and Schott, 2016). In this paper, we characterize welfare changes and labor reallocation in response to such sectoral shocks, taking into account the full general-equilibrium effect in a dynamic environment.

On the more structural side, recent papers have emphasized the importance of transitional dynamics in studying the effect of sectoral shocks and built models based on the dynamic discrete choice framework initially introduced in the IO literature (e.g., Rust, 1987). An important early contribution is Artuç, Chaudhuri, and McLaren (2010), who adopt this framework to analyze the impact of a trade shock on labor market dynamics. Subsequent papers enrich this framework by incorporating more realistic elements to investigate the effect of the China shock—including trade and input-output linkages (Caliendo, Dvorkin, and Parro, 2019); involuntary unemployment due to downward nominal wage rigidities or search frictions (Rodriguez-Clare, Ulate, and Vasquez, 2022); and endogenous trade imbalances (Dix-Carneiro et al., 2023). We extend the literature by introducing arbitrary time-invariant worker heterogeneity to the framework and show how this heterogeneity significantly affects the results of counterfactual exercises.

In allowing for persistent worker heterogeneity, we relate to a large, mostly static literature that emphasizes self-selection based on comparative advantage (see Borjas, 1987; Heckman and Sedlacek, 1985; and Ricardo's theory of comparative advantage for early contributions and Lagakos and Waugh, 2013; Young, 2014; Burstein, Morales, and Vogel, 2019; Hsieh et al., 2019; Porzio, Rossi, and Santangelo, 2022; Grigsby, 2022; and Adao, Beraja, and Pandalai-Nayar, 2023 for recent applications and developments). The work most closely related to ours includes Costinot and Vogel (2010); Adão (2016); Lee (2020), and Galle, Rodríguez-Clare, and Yi (2023), who also study the distributional effects of trade shocks. We contribute to this literature by embedding this mechanism within a dynamic discrete choice framework, which allows us to take transitional dynamics into account and highlight the importance of self-selection in a dynamic context.

In terms of methodology, we follow the sufficient statistics tradition (e.g., Chetty, 2009; Hulten, 1978; Arkolakis, Costinot, and Rodríguez-Clare, 2012; Baqaee and Farhi, 2020; McKay and Wolf, 2022; Beraja, 2023). In the present context, the use of sufficient statistics offers multiple advantages. First, by eliminating the need to estimate numerous primitives, it reduces computational costs and identification requirements while ensuring the transparency of the analysis. Second, our approach allows us to accommodate arbitrary worker heterogeneity without having to deal directly with the well-known identification challenges associated with unobserved heterogeneity (see, e.g., Heckman and Honore, 1990; French and Taber, 2011).

Finally, our paper is also related to a large structural literature that incorporates rich heterogeneity—such as age, gender, education, nonpecuniary benefit, tenure, unobserved comparative advantage—in dynamic models. Prominent examples include Keane and Wolpin (1997) and Lee and Wolpin (2006) in the labor literature and Dix-Carneiro (2014) and Traiberman (2019) in the trade literature. Our contribution to this body of literature is twofold. First, our results may serve to guide future structural analysis by suggesting that one can include worker flows over different time horizons as targeted moments in structural estimations to properly capture labor market dynamics. Second, our sufficient statistics provide a transparent way to assess which types of heterogeneity in the literature matter more for counterfactual predictions of the model.

Outline. The remainder of the paper is organized as follows. Section 2 presents a model of labor market dynamics with arbitrary time-invariant worker heterogeneity and derives our main sufficient statistics results. Section 3 analytically shows how ignoring persistent worker heterogeneity systematically affects the sufficient statistics and in turn biases counterfactual predictions regarding welfare and labor market dynamics. Section 4 measures our sufficient statistics using US panel data and estimates the bias in sufficient statistics due to ignoring persistent heterogeneity. Section 5 applies our sufficient statistics approach to study the impact of a hypothetical trade liberalization and the China shock, and Section 6 concludes. The Appendix contains theoretical restuls and proofs omitted in the main text.

2. Dynamic Discrete Choice Model with Persistent Worker Heterogeneity

In this section, we present a model that extends the dynamic discrete choice model of Artuç, Chaudhuri, and McLaren (2010) (hereafter, ACM) by allowing arbitrary time-invariant worker heterogeneity. We focus on workers' dynamic discrete choice over sectors, though the same framework can be applied to analyze their geographic location or occupation choices by a simple relabeling.[2] We first describe the individual worker's problem, whose solution characterizes the dynamics of welfare and sectoral labor supply at individual level. We then show how we can aggregate the dynamics to macro level to derive equations that can be used to compute counterfactual changes in aggregate welfare and sectoral labor supply in response to sectoral shocks. We conclude by demonstrating how to combine this result with the labor demand side of the model to study the general equilibrium effects of sectoral shocks.

2.1 Workers' Dynamic Discrete Choice Problem

Time is discrete and indexed by t. There are S sectors indexed by $i, j \in \mathcal{S} = \{1, \ldots, S\}$. There is a continuum of infinitely lived heterogeneous workers. We allow for arbitrary time-invariant heterogeneity of workers by assigning each worker a type $\omega \in \Omega$, which is drawn from an unknown distribution W over Ω. Importantly, the type ω may capture not only observable demographic and socioeconomic characteristics, such as gender and education, but also unobserved differences across workers. At each point in time, workers decide in which sector to work. A worker ι of type ω who is employed in sector i at period t chooses in which sector to work in period $t+1$ in order to maximize her continuation value. The value of this worker in period t can be recursively written as

$$V_{it}^{\omega}(\iota) = w_{it}^{\omega} + \max_{j \in \mathcal{S}} \{\beta \, \mathbb{E}_t \, V_{jt+1}^{\omega}(\iota) - C_{ij}^{\omega} + \rho^{\omega} \cdot \varepsilon_{jt}(\iota)\}, \tag{1}$$

where the type-specific instantaneous utility, w_{it}^{ω}, can capture both the wage and nonpecuniary benefits from sector i in period t. For expositional purposes, we will refer to w_{it}^{ω} as sectoral wages (though in some of our applications, they will be the logarithm of the real wages). The term $C_{ij}^{\omega} \geq 0$ captures the type-specific cost of switching from sector i to j. The idiosyncratic shock $\varepsilon_{jt}(\iota)$ is worker-specific and reflects nonpecuniary motives for workers to switch sectors.[3] The expectation operator, $\mathbb{E}_t$, is taken over realizations of future wages

[2] Dynamic location or occupation choices have also been widely studied in the literature: e.g., location choices (Kennan and Walker, 2011; Amior and Manning, 2018; Allen and Donaldson, 2020; Bilal and Rossi-Hansberg, 2021; Howard and Shao, 2023; Kleinman, Liu, and Redding, 2023), occupation choices (Lee, 2005; Artuç and McLaren, 2015; Traiberman, 2019)

[3] Idiosyncratic shocks result in gross flows that are an order of magnitude larger than net flows, a pattern that is consistent with the observed data.

and future idiosyncratic shocks, conditional on the information available in period t. The parameter ρ^ω governs the relative importance of idiosyncratic shocks in sector choice decisions and hence determines the elasticity of sectoral employment with respect to sectoral wages. We allow the sectoral wages and switching costs to vary arbitrarily across different types of workers, as can be seen from the superscripts ω in equation (1).

When $|\Omega| = 1$—which means that there are no persistent differences across workers—our model reduces to the canonical homogeneous-worker sector choice model of ACM or, more broadly, to the standard dynamic discrete choice model (e.g., Rust, 1987). In these models, workers are ex ante homogeneous, and any ex post heterogeneity arising from different realizations of idiosyncratic shocks persists for only a single period. In the rest of the paper, we use the term *worker heterogeneity* exclusively to refer to persistent worker heterogeneity across different types of workers. When $C^\omega_{ij} = 0$ and $\rho^\omega = 0$, on the other hand, the worker problem boils down to choosing the sector that offers the highest wage, so this model reduces to the general Roy model of self-selection (e.g., Heckman and Honore, 1990).

As is standard in dynamic discrete choice models in trade, IO, and labor (e.g., Rust, 1987; Aguirregabiria and Mira, 2010), we assume that the idiosyncratic shocks, $\varepsilon_{jt}(\cdot)$, are drawn from a type I extreme-value distribution independently across workers, sectors, and time periods (see, e.g., McFadden, 1973). Let ex ante value, v^ω_{it}, be the expected value derived from choosing sector i in period t for workers of type ω, taking the average over the realizations of idiosyncratic shocks $\{\varepsilon_{jt}(\cdot)\}_j$, and let the transition probability, F^ω_{ijt}, be the probability that workers of type ω in sector i in period t choose sector j in period $t+1$. Standard extreme-value algebra gives an analytical characterization of the ex ante value and transition probabilities:

$$v^\omega_{it} \equiv \mathbb{E}_\iota V^\omega_{it}(\iota) = w^\omega_{it} + \rho^\omega \ln \sum_{j\in\mathcal{S}} \left(\exp(\beta\,\mathbb{E}_t\, v^\omega_{jt+1}) / \exp(C^\omega_{ij})\right)^{1/\rho^\omega}, \tag{2}$$

$$F^\omega_{ijt} \equiv \Pr_t(s_{t+1} = j | s_t = i, \omega) = \frac{\left(\exp(\beta\,\mathbb{E}_t\, v^\omega_{jt+1}) / \exp(C^\omega_{ij})\right)^{1/\rho^\omega}}{\sum_{k\in\mathcal{S}} \left(\exp(\beta\,\mathbb{E}_t\, v^\omega_{kt+1}) / \exp(C^\omega_{ik})\right)^{1/\rho^\omega}}, \tag{3}$$

where the expectation operator $\mathbb{E}_\iota$ is taken over the realizations of $\{\varepsilon_{jt}(\iota)\}_j$. Equation (2) expresses the value of being in sector i as the sum of the current period's instantaneous utility and a nonlinear aggregation of next-period values, net of switching costs. Equation (3) suggests that, all else being equal, workers are more likely to choose sectors with higher values net of switching costs. Applying the law of large numbers for a continuum of workers, we can characterize the law of motion of their sectoral employment share from their transition probabilities,

$$\ell^\omega_{jt+1} = \sum_{i\in\mathcal{S}} F^\omega_{ijt} \ell^\omega_{it}. \tag{4}$$

We also define the backward transition probability,

$$B^\omega_{jit} \equiv \Pr_t(s_t = i | s_{t+1} = j, \omega) = \frac{\ell^\omega_{it} F^\omega_{ijt}}{\ell^\omega_{jt+1}},$$

which is the probability that a type ω worker in sector j in period $t+1$ came from sector i in period t. We define $S \times S$ matrices F^ω_t and B^ω_t, whose (m,n)-element is F^ω_{mnt} and B^ω_{mnt}, respectively. We refer to them as the (forward) transition matrix and backward transition matrix, respectively. Note that the rows of these matrices sum to one.

2.2 Welfare and Labor Dynamics at the Micro Level

The system of equations (2)–(4) fully characterizes the labor supply side of the model. That is, given the series of sectoral wages, $\{w_{it}^{\omega}\}$, we can solve for the series of sectoral employment, $\{\ell_{it}^{\omega}\}$, and sectoral values, $\{v_{it}^{\omega}\}$, from this system of equations. These are the two variables of interest in our counterfactual analysis. As a first step toward deriving a sufficient statistics result, we consider infinitesimal sectoral shocks $\{\mathrm{d}w_{it}^{\omega}\}$ and take first-order approximations of equations (2) and (4) around a steady state. A *steady state* is associated with time-invariant sectoral wages ($w_{it}^{\omega} = w_{i}^{\omega}$ for all t), where the type-specific value, transition probabilities, and sectoral labor supply remain constant over time. In line with our previous notation, we denote the steady-state forward and backward transition matrices as F^{ω} and B^{ω}, respectively. The following equations summarize the responses of the endogenous variables in terms of deviations from a steady state:

$$\mathrm{d}v_t^{\omega} = \mathrm{d}w_t^{\omega} + \beta F^{\omega}\,\mathbb{E}_t\,\mathrm{d}v_{t+1}^{\omega}, \tag{5}$$

$$\mathrm{d}\ln\ell_{t+1}^{\omega} = B^{\omega}\,\mathrm{d}\ln\ell_t^{\omega} + \tfrac{\beta}{\rho^{\omega}}(I - B^{\omega}F^{\omega})\,\mathbb{E}_t\,\mathrm{d}v_{t+1}^{\omega}, \tag{6}$$

where we use the vector notation,

$$\mathrm{d}v_t^{\omega} = \begin{pmatrix}\mathrm{d}v_{1t}^{\omega} & \cdots & \mathrm{d}v_{St}^{\omega}\end{pmatrix}^{\intercal},\ \mathrm{d}w_t^{\omega} = \begin{pmatrix}\mathrm{d}w_{1t}^{\omega} & \cdots & \mathrm{d}w_{St}^{\omega}\end{pmatrix}^{\intercal},\ \text{and } \mathrm{d}\ln\ell_t^{\omega} = \begin{pmatrix}\mathrm{d}\ln\ell_{1t}^{\omega} & \cdots & \mathrm{d}\ln\ell_{St}^{\omega}\end{pmatrix}^{\intercal}.$$

The algebra follows that of Kleinman, Liu, and Redding (2023), as described in Appendix B.

Equations (5) and (6) demonstrate that knowledge of the steady-state transition matrices is enough to characterize the effect of sectoral shocks. However, these equations cannot be confronted directly with data if worker type ω is unobservable. The key idea of this paper is that despite this unobservability, these type-specific equations can be aggregated into equations that solely involve observable aggregate variables, which can be used to construct counterfactuals.

In order to prepare the aggregation at the macro level, we solve these equations forward and backward to write the responses of sectoral values and sectoral employment as a function of the expected past and future wage changes $(\cdots, \mathrm{d}w_{t-1}^{\omega}, \mathrm{d}w_t^{\omega}, \mathrm{d}w_{t+1}^{\omega}, \cdots)$. Lemma 1 summarizes the result.

Lemma 1. *For a given sequence of changes in sectoral wages $\{\mathrm{d}w_t^{\omega}\}$, the changes in type-specific sectoral values and sectoral employment $\{\mathrm{d}v_t^{\omega}, \mathrm{d}\ln\ell_t^{\omega}\}$ are given by:*

$$\mathrm{d}v_t^{\omega} = \sum_{k\geq 0}(\beta F^{\omega})^k\,\mathbb{E}_t\,\mathrm{d}w_{t+k}^{\omega}, \tag{7}$$

$$\mathrm{d}\ln\ell_t^{\omega} = \tfrac{\beta}{\rho^{\omega}}\sum_{s\geq 0}(B^{\omega})^s(I - B^{\omega}F^{\omega})\left(\sum_{k\geq 0}(\beta F^{\omega})^k\,\mathbb{E}_{t-s-1}\,\mathrm{d}w_{t-s+k}^{\omega}\right). \tag{8}$$

Workers are forward-looking, so all future shocks affect the value of workers and sectoral employment. Due to the presence of switching costs and idiosyncratic shocks, labor reallocation is sluggish, so past shocks also affect current sectoral employment.

2.3 Welfare and Labor Dynamics at the Macro Level

We focus on the effect of sectoral shocks on sector-level variables aggregated across workers of different types ω. In particular, we define aggregate sectoral employment share and average sectoral value as follows:

$$\ell_{it} = \int_{\Omega} \ell_{it}^{\omega} \, \mathrm{d}W(\omega) \quad \text{and} \quad v_{it} = \int_{\Omega} v_{it}^{\omega} \, \mathrm{d}W(\omega|s=i) \tag{9}$$

where $W(\cdot|s=i)$ is the steady-state type distribution of workers in sector i. The total employment of sector i is obtained by summing the employment of different types of workers. Likewise, the average value of workers in sector i is given by taking the weighted average across different types of workers, using the steady-state type distribution of that sector as weights. In so doing, we implicitly assume a utilitarian social welfare function with equal weights across all workers.

Next, we define worker flow matrices.

Definition 1. *For each $k \in \mathbb{N}_0$, the k-*period worker flow matrix *$\mathcal{F}_k$ is an $S \times S$ matrix whose (i,j)-element is given by the steady-state population share of workers in sector i who switch to sector j after k periods:*

$$(\mathcal{F}_k)_{i,j} = \Pr(s_{t+k} = j | s_t = i).$$

Unlike the type-specific transition matrices B^{ω} and F^{ω} in Lemma 1, these worker flow matrices can be computed directly from longitudinal information on workers' sector choices, as we will do with the NLSY data in Section 4. As we will demonstrate, these matrices are sufficient statistics for characterizing the welfare and labor market consequences of sectoral shocks.

To derive an aggregation result, we make the following two assumptions.

Assumption 1. *Workers of different types share common sectoral shocks and common dispersion of idiosyncratic shocks:*

$$\mathrm{d}w_t^{\omega} = \mathrm{d}w_t \;\text{ and }\; \rho^{\omega} = \rho, \textit{ for all } t \textit{ and } \omega \in \Omega.$$

Assumption 2. *The bilateral switching costs between sectors are* quasi-symmetric*; that is, they can be decomposed into*

$$C_{ij}^{\omega} = C_i^{\omega} + \tilde{C}_j^{\omega} + \hat{C}_{ij}^{\omega}$$

where $\hat{C}_{ij}^{\omega} = \hat{C}_{ji}^{\omega}$ for all $i, j \in \mathcal{S}$, and $\omega \in \Omega$.

It is worth emphasizing that the first part of Assumption 1 allows that workers have different *level* of instantaneous utilities. For example, suppose workers have log utility and shocks are multiplicative to the wages of all workers. In this case, even if workers have different wages, the shocks manifest themselves as common additive shocks to instantaneous utilities for all workers. In addition, we can account for the possibility that shocks have heterogeneous effects across workers with different *observable* characteristics—for example, between high-skilled and low-skilled workers or across geographic regions—by conducting the same analysis separately for each characteristic.[4] Similarly, the second part of Assumption 1 allows that workers have different labor supply elasticities. Although the dispersion of idiosyncratic shocks governs the elasticity of sectoral labor supply with respect to sectoral shocks, the elasticity also depends on type-specific transition matrices, B^{ω} and F^{ω}, as can be seen in equation (8). Although restrictive, this assumption is standard in the dynamic discrete choice literature, even in papers that incorporate rich heterogeneity of workers (e.g., Keane and Wolpin, 1997; Lee and Wolpin, 2006; Arcidiacono and Miller, 2011; Dix-Carneiro, 2014; Traiberman, 2019).

In the literature, quasi-symmetry in Assumption 2 and its special cases are often imposed on bilateral switching costs and bilateral trade costs for various purposes. For example, many papers adopt this assumption in order to reduce the number of parameters to be estimated (Eaton and Kortum, 2002; Artuç, Chaudhuri, and McLaren, 2010; Dix-Carneiro, 2014). Alternatively, other works, such as Allen and Arkolakis (2014), Desmet, Nagy, and Rossi-Hansberg (2018), and Allen, Arkolakis, and Takahashi (2020), impose this condition to simplify the equilibrium system into a single integral equation.

We are ready to state our main result.

Proposition 1. *Suppose that Assumptions 1 and 2 hold. For a given sequence of (common) changes in sectoral wages $\{\mathrm{d}w_t\}$, the changes in sectoral value and sectoral employment are given by:*

$$\mathrm{d}v_t = \sum_{k\geq 0} \beta^k \mathcal{F}_k \, \mathbb{E}_t \, \mathrm{d}w_{t+k}, \tag{10}$$

$$\mathrm{d}\ln \ell_t = \sum_{s\geq 0, k\geq 0} \frac{\beta^{k+1}}{\rho} (\mathcal{F}_{s+k} - \mathcal{F}_{s+k+2}) \, \mathbb{E}_{t-s-1} \, \mathrm{d}w_{t-s+k}. \tag{11}$$

The logic behind this proposition is as follows. Starting from Lemma 1, we want to aggregate type-specific variables to the macro level. We first invoke Assumption 2, which simplifies the aggregation by giving the equality between the forward and backward transition matrices, $B^{\omega} = F^{\omega}$.[5] Although this equality is not strictly necessary for our purpose, it allows us to derive analytical results in Section 3 and reduces the data requirements needed to implement our sufficient statistics approach.[6] These two transition matrices are indeed very similar at the level of granularity at which both matrices can be computed from the data; see Figure OA.6 in the online appendix. After imposing this equality, we aggregate equations (7) and (8) to derive equations (10) and (11), respectively. In particular, the kth powers of the type-specific transition matrix $(F^{\omega})^k$ are aggregated into the k-period worker flow matrix $\mathcal{F}_k$. Intuitively, since the (i, j)-element of the former is given by $\Pr(s_{t+k} = j | s_t = i, \omega)$, we can obtain the (i, j)-element of the latter, $\Pr(s_{t+k} = j | s_t = i)$, by

[4] This is the approach we take when we study the effect of the China shock on US labor markets in Section 5. However, this approach is infeasible for *unobserved* types. Therefore, we cannot dispense with the assumption that shocks are common across unobserved types.

[5] See Appendix A.2 for a proof. Two matrices are equal if and only if the steady-state bilateral worker flows are symmetric. As in Allen, Arkolakis, and Takahashi (2020), quasi-symmetry in Assumption 2 is sufficient to guarantee it.

[6] See Appendix A.2 for a general version of Proposition 1 without this assumption.

taking an average over the type distribution of workers in sector i, $\Pr(\omega|s_t = i)$. In Appendix A.1, we prove Proposition 1 by introducing the population-average operator, which formalizes the idea of aggregation.[7]

The role of Lemma 1 and Assumption 1 should also be clear at this point. We have just seen that products of transition matrices can be aggregated to worker flow matrices, but when they are multiplied by another type-specific variable, such as $\mathrm{d}v^{\omega}_{t+1}$ in (5) and $\mathrm{d}w^{\omega}_{t+k}$ in (7), a complication arises because the aggregation then involves a covariance term that captures the extent to which two multiplicands comove across different worker types. Since the type index ω may include unobserved heterogeneity, it is not possible to characterize the covariance term without specifying the precise form of worker heterogeneity. Lemma 1 and Assumption 1 allow us to bypass this problem.

Proposition 1 establishes that in order to calculate the counterfactual changes in aggregate welfare and sectoral employment for a known sequence of exogenous wage changes, $\{\mathrm{d}w_t\}$, we only require knowledge of the worker flow matrices, $\{\mathcal{F}_k\}$, and the two parameters, ρ and β. In particular, we do not need full knowledge of the detailed worker heterogeneity (i.e., the distribution of types, W) and resulting self-selection that generates these worker flow matrices. What matters is how frequently workers switch sectors over time, not the specific structural determinants of these patterns.[8] In Section 4, we estimate worker flow matrices and the value of the parameter ρ using panel data. The following corollary summarizes the discussion.[9]

Corollary 1. *Consider a sequence of exogenous changes in sectoral wages, $\{\mathrm{d}w_t\}$. Along with ρ and β, the worker flow matrices, $\{\mathcal{F}_k\}$, constitute sufficient statistics for changes in sectoral values, $\{\mathrm{d}v_t\}$, and sectoral employment, $\{\mathrm{d}\ln \ell_t\}$.*

The worker flow matrix $\mathcal{F}_k$ is informative about the welfare effect because it captures workers' k-period sector choices. It is also informative about the response of sectoral employment shares, as it reveals how many workers are at the margin between each pair of sectors. A more detailed intuition, along with the role of the extreme-value distribution assumption, is provided in Appendix A.4.

At this point, it is worth discussing how our sufficient statistics approach relates to structural work in this area. Our approach stands in stark contrast to the standard structural approach to accounting for worker heterogeneity. First, our approach eliminates the need to estimate many primitives. This reduces computational costs and data requirements, while ensuring the transparency of the analysis. Like other studies of sufficient statistics (see, e.g., Chetty, 2009), our method yields counterfactual predictions that are immune to the Lucas critique, without requiring knowledge of the full structure of the model. Second, this

[7] In static settings, a similar aggregation idea for the first-order welfare effect is advanced in Kim and Vogel (2020) and Sprung-Keyser, Hendren, and Porter (2022), under conditions similar to the first part of Assumption 1.

[8] This result is surprising because, in principle, constructing the counterfactual in a dynamic context without a precisely specified model requires estimating all dynamic elasticities of sectoral values and sectoral employment with respect to shocks at all time horizons—that is, how past as well as future shocks affect these variables. McKay and Wolf (2022) propose a method to operationalize this approach in practice, but in general estimating all elasticities is challenging due to high information requirements and limited availability of data. Proposition 1 reveals that the dynamic discrete choice framework imposes a tight connection among these dynamic elasticities. This relationship enables us to parameterize these elasticities with a single parameter to be estimated, ρ, while the worker flow matrices contain all the remaining information needed to calculate the dynamic elasticities.

[9] In Arkolakis, Costinot, and Rodríguez-Clare (2012), a distinction is made between the ex ante sufficient statistics result and the ex post result. Proposition 1 provides the ex post sufficient statistics in the sense that this result is only useful if we can estimate or directly observe the change in wages resulting from the shock of interest. This is often feasible when examining the effects of shocks that occurred in the past. However, it becomes impossible when attempting to forecast the impact of hypothetical shocks.

approach effectively accommodates arbitrary worker heterogeneity without encountering the well-known challenges associated with estimating the distribution of unobserved heterogeneity from the data.[10]

These advantages, however, do not come without costs. In addition to the Assumptions 1 and 2, we rely on a first-order approximation around a steady state, the validity of which depends on the sectoral shock of interest. Nonetheless, as shown by Kleinman, Liu, and Redding (2023), even when the economy is not initially in steady state, the transition dynamics of sectoral value and employment shares can be additively decomposed into the contributions of sectoral shocks and the convergence dynamics toward the steady state. This result implies that it is without loss of generality to assume that the economy is initially in steady state when computing the effect of sectoral shocks. Furthermore, in Section 5, we demonstrate that the first-order approximation performs well for the magnitude of shocks we consider.

2.4 Closing the Model: Labor Demand and Equilibrium Wages

We have so far focused on the labor supply side and considered exogenously given wage changes. We conclude this section by extending the sufficient statistics result to the case in which the wage is endogenously determined by the labor market equilibrium. As it turns out, even when we endogenize the wage, the same set of sufficient statistics, combined with knowledge of the labor demand side, constitutes sufficient statistics for counterfactual changes in sectoral values and sectoral employment.

Although the main contribution of this paper centers on the labor supply side, our heterogeneous worker labor supply model can be integrated with any labor demand system. The specific nature of the labor demand system depends on preferences, technology, and good market structure. For expositional purposes, we specify it in a reduced-form manner in this section and return to its structural determinants in our applications in Section 5. Specifically, we assume that the wage of sector i is endogenously determined by the sectoral labor allocation $\{\ell_{jt}\}_j$ and exogenous shocks $\{\xi_{jt}\}_j$ that affect the marginal productivity of labor. The variable ξ_{jt} encompasses sector-specific factors, such as capital stock, technology shocks, policy variables, and the like. This relationship can be expressed as $w^{\omega}_{it} = f^{\omega}_i(\{\ell_{jt}\}_j, \{\xi_{jt}\}_j)$. In Appendix A.3, we show that under Assumption 1 we can write this relationship in terms of a first-order approximation as

$$\mathrm{d}w_t = D \cdot \mathrm{d}\ln \ell_t + E \cdot \mathrm{d}\xi_t, \tag{12}$$

where there is no ω-index on matrices D and E.

Combining the labor demand curve represented by this equation with the labor supply curve characterized in the previous proposition, we can define the labor market equilibrium. It consists of paths of type-specific sectoral value, v^{ω}_t; type-specific labor allocation across sectors, ℓ^{ω}_{t+1}; type-specific transition probabilities, F^{ω}_t; aggregate sectoral value, v_t; and aggregate labor allocation, ℓ_{t+1}, that are measurable with respect to the period-t information set, and a path of sectoral wages, w_t, that are measurable with respect to the period-$(t-1)$ information set and the period-t shock such that: *(a)* type-specific variables $\{v^{\omega}_t, \ell^{\omega}_{t+1}, F^{\omega}_t\}$

[10] Keane and Wolpin (1997); Lee and Wolpin (2006); Dix-Carneiro (2014) and Traiberman (2019) incorporate unobserved worker types in their analyses. However, due to the identification challenge, they are limited to using a small number of unobserved types—between two and four. While this approach can capture workers' absolute advantages, it is difficult to fully capture the comparative advantage of workers and hence their self-selection into different sectors.

solve problem (1) given the path of wages; *(b)* aggregate variables $\{v_t, \ell_t\}$ are consistent with the type-specific variables through equation (9); *(c)* wages are determined by the marginal productivity of labor, (12); and *(d)* the labor market clears.

The following proposition shows that the same set of worker flow matrices, combined with knowledge of the labor demand side, still constitutes sufficient statistics when wages are endogenously determined by the labor market equilibrium.

Proposition 2. *Suppose that Assumptions 1 and 2 hold. For a given sequence of labor demand shocks $\{\mathrm{d}\xi_t\}$, the equilibrium values of $\{\mathrm{d}v_t, \mathrm{d}\ln\ell_t, \mathrm{d}w_t\}_t$ are given by solution of the following system of equations:*

$$\begin{aligned}
\mathrm{d}v_t &= \sum_{k\geq 0} \beta^k \mathcal{F}_k \, \mathbb{E}_t \, \mathrm{d}w_{t+k}, \\
\mathrm{d}\ln\ell_t &= \sum_{s\geq 0, k\geq 0} \frac{\beta^{k+1}}{\rho} (\mathcal{F}_{s+k} - \mathcal{F}_{s+k+2}) \, \mathbb{E}_{t-s-1} \, \mathrm{d}w_{t-s+k}, \\
\mathrm{d}w_t &= D \cdot \mathrm{d}\ln\ell_t + E \cdot \mathrm{d}\xi_t.
\end{aligned}$$

The intuition is simple. Conditional on a path of wage changes across time and across sectors, we can characterize the dynamic response of sectoral employment using Proposition 1. Conditional on the dynamics of sectoral employment, we can solve for the path of wage changes from the labor demand side. The equilibrium is determined as a fixed point of these relations.

Unlike Proposition 1, which requires knowledge of the path of wages, Proposition 2 requires the path of labor demand shocks $\{\mathrm{d}\xi_t\}$. Given the shock path, $\{\mathrm{d}\xi_t\}$, we can solve the system of equations to compute changes in values and sectoral employment. Conditional on the observed series of worker flow matrices, the values of ρ and β, and the reduced-form specification of the labor demand side, D and E, the responses of welfare, employment, and wages to labor demand shocks do not depend on the specific details of worker heterogeneity.

3. Employment and Welfare Implications of Persistent Heterogeneity

The canonical dynamic discrete choice model commonly used in the literature abstracts from persistent worker heterogeneity. The sufficient statistics results in the previous section provide a way to account for arbitrary time-invariant heterogeneity. In this section, we use these results to demonstrate why worker heterogeneity matters when evaluating the consequences of sectoral shocks.

Our sufficient statistics result highlights that worker heterogeneity affects the results of counterfactual exercises only through its effect on the model's predictions for a particular set of moments of the data: the worker flow matrices $\{\mathcal{F}_k\}$. In this section, we first theoretically characterize a systematic bias in worker flow matrices implied by the canonical model, which, due to the lack of persistent worker heterogeneity, imposes that the k-period worker flow matrix is equal to the one-period worker flow matrix to the kth power. In turn, the bias in worker flows leads to systematic biases in counterfactual predictions of welfare changes and labor reallocation.

3.1 Steady-state Worker Flow with and without Persistent Heterogeneity

Lemma 2 characterizes the restrictions that the canonical model imposes on the worker flow matrices $\{\mathcal{F}_k\}$, and illustrates how accommodating worker heterogeneity relaxes this restriction. A formal proof is provided for a special case of our model with two sectors; however, the same inequality holds in the data for worker flow matrices involving more than two sectors (see Section 4.2).

Lemma 2. *Without persistent worker heterogeneity ($|\Omega| = 1$), we have $\mathcal{F}_k = (\mathcal{F}_1)^k$. Suppose $S = 2$. With (non-degenerate) persistent worker heterogeneity, we have $(\mathcal{F}_k)_{ii} > \left((\mathcal{F}_1)^k\right)_{ii}$ for all $i \in \mathcal{S}$ and $k > 1$.*

Without persistent heterogeneity, the Markovian structure of the model implies that the same transition probabilities apply to all workers, which enables us to compute the k-period worker flow matrix by multiplying the one-period matrix k times. With worker heterogeneity, however, the diagonal elements of the k-period worker flow matrices could be larger than they would be in the absence of worker heterogeneity. Accordingly, if we wrongly ignore persistent worker heterogeneity, we systematically *underestimate* the probability of workers choosing the same sector after k periods, concluding that moving across sectors is more frequent than it actually is. To understand this underestimation, consider the case of $k = 2$, where we have

$$(\mathcal{F}_2)_{ii} = \sum_{j \in \mathcal{S}} \Pr(s_{t+1} = j | s_t = i) \Pr(s_{t+2} = i | \boldsymbol{s_t = i}, s_{t+1} = j), \tag{13}$$

$$\left((\mathcal{F}_1)^2\right)_{ii} = \sum_{j \in \mathcal{S}} \Pr(s_{t+1} = j | s_t = i) \Pr(s_{t+2} = i | s_{t+1} = j).$$

When workers are heterogeneous, the additional conditioning of $s_t = i$ in equation (13) increases the likelihood of choosing sector i again in period $t+2$, since workers who have self-selected into sector i in period t are more likely to do so in subsequent periods. This result is reminiscent of the findings in Heckman (1981) and Heckman and Singer (1984) that the average hazard rate is biased toward negative duration dependence relative to each type-specific hazard rate.

Many widely used datasets only provide information on the short-run worker flows because they do not track individual workers, nor do they provide information on workers' past sector choice history or tenure.[11] In such situations, a common approach in the literature is to assume homogeneous workers and calibrate models by matching the one-period worker flow matrix. Lemma 2 shows how this calibration practice effectively extrapolates longer-run worker flows and why this extrapolation is necessarily biased. In Section 4, we indeed show that the canonical model performs poorly in matching the longer-run worker flow patterns observed in the data.

3.2 Counterfactual Predictions with and without Persistent Heterogeneity

Combining Lemma 2 with our sufficient statistics result, we can theoretically characterize the systematic biases in counterfactual predictions that arise from assuming away persistent heterogeneity. For the moment,

[11] Even when researchers use panel data that contain the necessary information, it is unclear whether the model correctly matches longer-run worker flows—or if it can match them at all.

we consider shocks to exogenously given wages, as in Proposition 1. For simplicity, we focus on a uniform permanent shock, either positive or negative, to a sector $s \in \mathcal{S}$ that is known to workers in period 1:

$$\mathrm{d}w_{st} = \Delta \in \mathbb{R}, \ \ \forall t \geq 1. \tag{14}$$

For more general shocks, we characterize the effect of one-time shocks, from which we can calculate the effect of any sequence of shocks; see Section OA.1.2 in the online appendix.

Counterfactual Welfare Changes. We begin with welfare changes. The inequality in Section Lemma 2 implies that, compared to the predictions of the canonical model, workers who are initially employed in sector s are more likely to remain in the sector and to be affected by the wage change for a longer period of time. As a result, ignoring worker heterogeneity leads to underestimation of the welfare changes of these workers. This observation is formalized in Proposition 3.

Proposition 3. *Consider a uniform permanent shock of the form* (14) *known to workers in period* 1. *The inequality in Lemma 2 implies that, the canonical model, calibrated by matching the one-period worker flow matrix, underestimates the welfare effect on workers initially employed in sector s,* $|\,\mathrm{d}v_{s1}|$.

Counterfactual Employment Changes. We now turn to labor reallocation. A shock to sector s changes the employment share of that sector over time. This labor reallocation is characterized by equation (11) of Proposition 1, which involves terms of the form $\mathcal{F}_k - \mathcal{F}_{k+2}$. For ease of notation, we define b_k as the diagonal element of $\mathcal{F}_k - \mathcal{F}_{k+2}$ that corresponds to sector s:

$$b_k \equiv (\mathcal{F}_k - \mathcal{F}_{k+2})_{s,s}.$$

Roughly speaking, b_k measures the rate at which the probability of remaining in sector s declines over time. To characterize the bias of the canonical model, we assume a single-crossing condition on b_k.

Assumption 3. *There exists $\bar{k} \in \mathbb{N}$ such that b_k is higher in the canonical model if and only if $k \leq \bar{k}$.*

This assumption requires that the probability of remaining in sector s initially decreases faster in the canonical model, but eventually decreases faster in the heterogeneous-worker model. Note that both models give the same value of the staying probabilities $(\mathcal{F}_k)_{s,s}$ for $k = 0, 1$, and the heterogeneous-worker model yields higher staying probabilities for all $k \geq 2$. Thus, staying probabilities must decrease faster in the canonical model for early periods. On the other hand, if staying probabilities converge to similar levels in both models as $k \to \infty$, then the decline should eventually become faster in the heterogeneous-worker model in order to compensate for the initial slower decline. Assumption 3 further requires the existence of a cutoff $\bar{k}$ at which the order of the speed of decline is reversed. In Section OA.1.1, we show that this assumption indeed holds with $\bar{k} = 9$ (years) for the worker flow matrix series we observe in the data. Under this assumption *on the steady-state worker flows*, the following proposition shows that the canonical model initially overestimates the change in employment in sector s while underestimating the long-run labor reallocation *in response to the shock*.

Proposition 4. *Consider a uniform permanent shock of the form* (14) *known to agents in period* 1*. Under Assumption 3, there exists* $\bar{t} \in \mathbb{N} \cup \{\infty\} \backslash \{1\}$ *such that the canonical model, calibrated by matching the one-period worker flow matrix, overestimates the change in employment of sector s in period t if and only if* $1 < t \leq \bar{t}$.

The result implies that whether assuming away persistent heterogeneity leads to overestimation or underestimation of the labor reallocation depends on the time horizon. On the one hand, as discussed in Lemma 2, the canonical model overestimates the mobility of workers across sectors, leading to an overestimation of the change in employment of sector s. This intuition is what Proposition 4 describes when t is small. On the other hand, in the heterogeneous-worker model, once workers choose sector s, they have relatively higher probabilities of being stuck in that sector. Thus, in the face of a permanent negative (positive, respectively) shock, workers will dislike (like, respectively) sector s more relative to the canonical model. This aspect works in the opposite direction to our previous intuition and may become dominant when t is large enough.[12] In Appendix A, we show that the worker flow matrix series we observe in the data implies $\bar{t} = 11$ years. Thus, we can conclude that the canonical model overestimates the impact of shocks on sectoral employment within an 11-year horizon but underestimates their longer-run effects.

Until now, we have considered exogenous changes in sectoral wages. However, if wage changes are endogenously determined by labor market equilibrium, different models may also generate different predictions for wage changes in response to given exogenous shocks to the labor market. Interestingly, with endogenously determined wages, the underestimation of the welfare effect characterized in Proposition 3 is likely to be compounded by the overestimation of the speed of labor reallocation shown in Proposition 4. To see this, consider a negative shock to a sector. Proposition 4 implies that in response to a given negative wage change, workers leave the sector more rapidly in the canonical model. The resulting decrease in labor supply raises the marginal productivity of labor in that sector, which partially offsets the initial decline in wages. Thus, the canonical model predicts a smaller decline in wages, at least in the short term. This, combined with discounting of the future, further contributes to the underestimation of the welfare effect.

In sum, we show that the canonical model, without persistent worker heterogeneity, underestimates the welfare losses of adversely affected workers and overestimates the speed of labor reallocation. In our counterfactual exercises in Section 5, we indeed document sizable differences in welfare effects and labor reallocation with and without persistent worker heterogeneity.

4. Sufficient Statistics in the Data

The sufficient statistics result in Proposition 2 requires three sets of inputs to construct counterfactuals for a given shock of interest: the worker flow matrix series, values of the parameters ρ and β, and knowledge of the labor demand side. In this section, we first use longitudinal information in the NLSY data to compute the aggregate worker flow matrices and compare them with those implied by the canonical model without persistent worker heterogeneity. We then present a method for estimating the value of ρ without specifying

[12] This is not always the case. In particular, there exists $\bar{\beta} \in (0, 1)$ such that when $\beta > \bar{\beta}$, the canonical model overestimates the change in employment in sector s in all periods t.

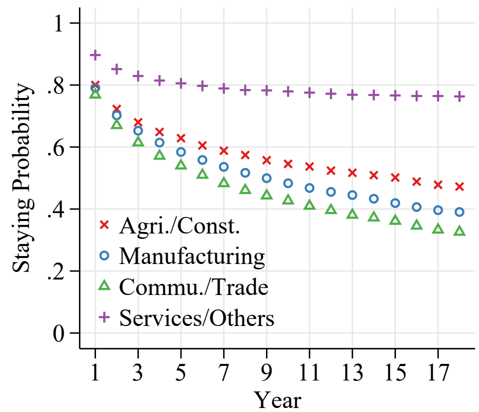


Figure 2. Worker Flow Matrix Series

Notes: Each marker in the figure represents the probability that workers choose the same sector after $k \in \{1, 2, \ldots, 18\}$ years, $\Pr(s_{t+k} = s | s_t = s)$. There are four sectors: Agriculture and Construction, Manufacturing, Communication and Trade, and Services and Others. Data source: NLSY79.

worker heterogeneity, which extends the standard Euler-equation approach used in the literature. Finally, we impose $\beta = 0.96$ for the subsequent analysis. In Section 5, we close the model by specifying details on the labor demand side for each application.

4.1 Observed Worker Flow Matrices

We compute worker flow matrices from the National Longitudinal Survey of Youth 1979, a rich dataset compiled by the US Bureau of Labor Statistics.[13] This survey follows a nationally representative sample of workers from 1979 onward annually through 1994 and biennially thereafter. The sample consists of workers who were between 14 and 21 years old as of December 31, 1978, and entered the labor market in the 1980s. The NLSY79 provides detailed information on education, race, gender, age, and, importantly, the sector of employment. Specifically, we identify a worker's sector of employment in a year as the sector in which the worker was employed in the first week of that year, which ensures that we consistently measure mobility at a 1-year window. We mainly follow the data-cleaning procedure of Lise and Postel-Vinay (2020).[14] We consider worker flows across four broadly defined sectors: (i) Agriculture and Construction; (ii) Manufacturing; (iii) Communication and Trade; and (iv) Services and Others.

Assuming that the economy was in a steady state between 1980 and 2000, we calculate the series of worker flow matrices by pooling all observations in this period.[15] Given the data constraints, we only calculate the k-year worker flow matrices, $\mathcal{F}_k$, up to $k = 18$. Figure 2 plots the diagonal elements of the obtained

[13] We also obtain quantitatively and qualitatively similar findings using the monthly Current Population Survey dataset; see Figure OA.15.

[14] The survey comprises a cross-sectional subsample representative of young people living in the US and other subsamples that target ethnic minorities, people in the military, and the poor. We only use the representative subsample for our analysis. We also drop people seen in the military.

[15] Figure A.2 shows how the sectoral employment shares change over time during this period. The shares remain stable, although the employment share of Services/Others exhibits a slow increase.

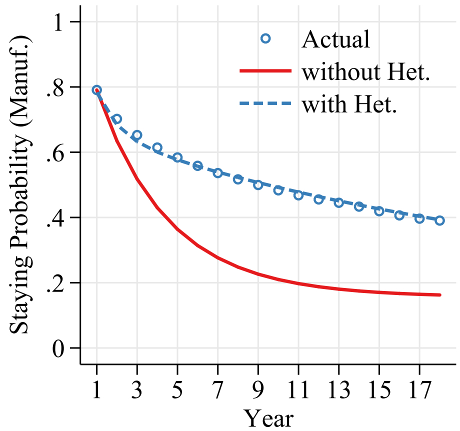


Figure 3. Actual and Model-implied Worker Flow Matrices: Manufacturing Staying Prob.

Notes: Blue dots in the figure represent k-year manufacturing staying probabilities. The blue solid line represents the fit of the estimated two-type worker model. The red solid line represents staying probabilities implied by the canonical model. Data source: NLSY79.

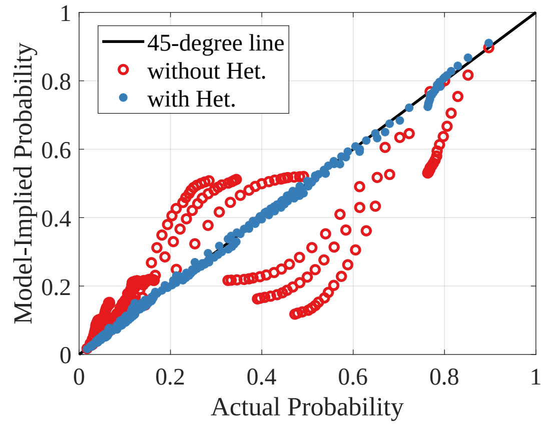


Figure 4. Actual and Model-implied Worker Flow Matrices: All 4×4 Elements

Notes: Blue dots in the figure plot elements of the worker flow matrix series implied by the estimated two-type worker model against those in the data. Red dots correspond to the canonical model. Data source: NLSY79.

worker flow matrices. Each point represents the probability of workers choosing the same sector after k years (i.e., $\Pr(s_{t+k} = s|s_t = s)$; hereafter, *k-year staying probability*). 1-year staying probabilities are close to 80%, except in the services sector, in which it is just below 90%. k-period staying probabilities decrease in k, which reflects the diminishing impact of being in a particular sector in the past over time.

4.2 Bias of the Canonical Model

To quantify the bias of the canonical model characterized in Lemma 2, we calculate the worker flow matrix series implied by the canonical model. Following Lemma 2, we compute the implied k-year worker flow matrix by multiplying the 1-year matrix k times. We first compare a specific diagonal element of the actual and the implied worker flow matrices: the k-year staying probability for the manufacturing sector. This is the element of primary interest because our main counterfactual exercise in Section 5 examines the impact of the China shock on US manufacturing sectors. Figure 3 plots both the actual staying probabilities (blue dots) and those implied by the canonical model (red line). It clearly shows that the canonical model significantly underestimates longer-run staying probabilities, which is in line with the prediction of Lemma 2 and becomes particularly pronounced at longer horizons.[16] Figure A.1 further shows that this underestimation is not driven

[16] One concern is that this bias may mainly, or at least partly, reflect misclassification errors in the sector choice data. Dvorkin (2023) studies the bias resulting from misclassification errors in the industry (and occupation) information in the PSID and March CPS datasets. Such misclassification errors lead to overestimation of mobility, as in this paper. However, it is unlikely that our results are mainly due to these errors. First, the NLSY data and the monthly CPS (observations within four consecutive months) used in this paper are known to be less prone to such errors (Moscarini and Thomson, 2007). Second, a similar pattern has been documented in other contexts. For example, Howard and Shao (2023) find it in the context of dynamic location choices, which is less prone to to misclassification errors.

by the nonstationary nature of the data. The result remains qualitatively and quantitatively similar even when nonstationarity is taken into account.[17]

As we have seen in Section 3, this discrepancy arises because the likelihood of choosing the manufacturing sector is higher for workers who have previously chosen the manufacturing sector. Workers who have self-selected to stay in manufacturing exhibit a higher probability of staying again in the following year, perhaps due to particularly high switching costs. Similarly, workers who self-selected into manufacturing in the past are more likely to choose it again, possibly owing to their comparative advantage.

The canonical model also underestimates the diagonal elements of the worker flow matrices that correspond to the non-manufacturing sectors. In Figure 4, we plot all 4×4 elements of the worker flow matrix series implied by the canonical model against the actual values in the data (red dots). The points clustered below the 45-degree line correspond to the diagonal elements of the worker flow matrices, which are underestimated by the canonical model. To compensate for this underestimation, the off-diagonal elements are overestimated, as seen in other points clustered above the 45-degree line. Section 5 quantifies how this inconsistency translates into systematic biases in counterfactual predictions of the effects of sectoral shocks.

4.3 Understanding the Bias of the Canonical Model

Demographic and Socioeconomic Characteristics. Where does the bias of the canonical model come from? One possibility is that worker heterogeneity in terms of observable characteristics could explain most of the bias. If so, we can easily correct for the bias by simply conditioning on these characteristics, obviating the need for our sufficient statistics approach. However, we will demonstrate that this is not at all the case. The literature has discussed various types of demographic and socioeconomic characteristics of workers. Here, we focus on four dimensions—gender, race, education, and age—that have been found to be important determinants of sectoral choices and welfare outcomes (e.g., Dix-Carneiro, 2014; Lee and Wolpin, 2006). We divided people into male and female; Hispanic/Black and non-Hispanic/Black; low-skilled (less than high school and high school) and high-skilled (some college and college or more); and young and old.[18] Unique combinations of these three dimensions of heterogeneity define sixteen worker types. Workers who differ along these characteristics exhibit highly distinct sectoral movement patterns; see Figure OA.7. For example, low-skilled non-Hispanic/Black males are more than twice as likely to stay in the manufacturing sector in the longer run than high-skilled Hispanic/Black females. However, Figure 5 shows that these characteristics do not explain the gap observed in Figure 3. We plot the manufacturing staying probabilities implied by the model that incorporates these four dimensions of observed characteristics. Specifically, we consider a model

[17] In fact, the stationarity assumption is likely to lead to an *underestimation* of the gap between the actual and implied staying probabilities. To see this, suppose that the worker flow matrix is F^1 in period $t = 1$ and F^2 in period $t = 2$. Under the stationarity assumption, we would calculate the worker flow matrix $\bar{F}$, which is applied to both periods, by taking an average of F^1 and F^2. Suppose that we put equal weights on F_1 and F_2. Then, we have

$$(F^1F^2)_{ii} \approx (F^1)_{ii}(F^2)_{ii} \leq \left(\frac{(F^1)_{ii} + (F^2)_{ii}}{2}\right)^2 = (\bar{F})^2_{ii} \approx (\bar{F}^2)_{ii}.$$

Thus, the implied two-period staying probability is overestimated under the stationarity assumption, leading to a *smaller* gap between the actual and implied staying probabilities.

[18] Age is a form of time-varying heterogeneity. To analyze this within the context of our persistent heterogeneity framework, we categorize age into two groups, "young" and "old," so that each represents about 50% of the total sample.

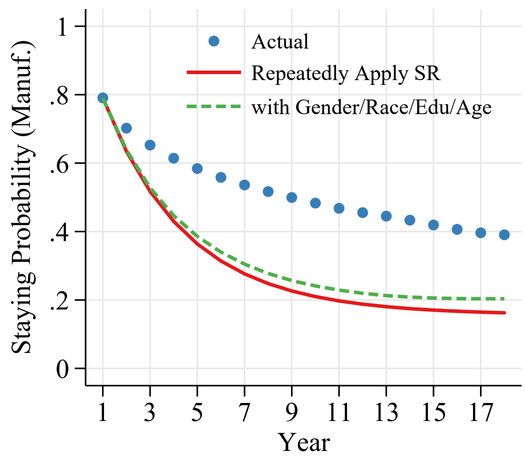


Figure 5. Actual and Model-implied Staying Probabilities, with and without Obs. Char.

Notes: Blue dots in the figure represent k-period manufacturing staying probabilities. The red solid line represents the fit of the estimated canonical model. The green dashed line incorporates four observed characteristics. Data source: NLSY79.

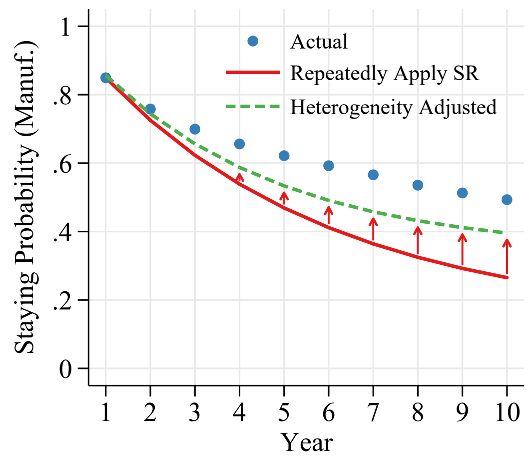


Figure 6. Actual and Model-implied Staying Probabilities, Heterogeneity Partially Controlled

Notes: Blue dots in the figure represent k-period manufacturing staying probabilities calculated using post-1990 data. Green dashed line and red line represent the staying probabilities implied by the canonical model and the model with five worker types, respectively. Data source: NLSY79.

with the sixteen observed worker types. For each worker type ω^{obs}, we can calculate the 1-year transition matrix $(F^{\omega^{\text{obs}}})$ directly from the data. Since workers are assumed to be homogeneous within each of the sixteen types, their k-year transition matrix can then be computed by $(F^{\omega^{\text{obs}}})^k$. Thus, the model-implied aggregate k-year worker flow matrix is obtained by taking averages of these type-specific k-year transition matrices using the steady-state type composition as the weight. The green dashed line shows the result and is almost indistinguishable from the red line, which plots the staying probabilities implied by the canonical model. These types of observed characteristics provide only a minor improvement in the model's ability to explain the actual pattern of worker flows. In sum, workers with different demographic and socioeconomic characteristics do indeed behave differently, but this fact barely changes *aggregate* labor market dynamics. In turn, the sufficient statistics result implies that adding these worker characteristics to a model does not cause substantial changes in the results of the counterfactual analysis.

Pure Duration Dependence. One concern might be that this gap does not reflect worker heterogeneity but is driven by other mechanisms such as pure duration dependence.[19] A notable example is the accumulation of sector-specific human capital, whereby otherwise identical workers who have spent more time in manufacturing may have accumulated more manufacturing-specific human capital, which renders them more likely

[19] Distinguishing dynamic selection based on heterogeneity from duration dependence mechanisms is a recurring theme in various fields of economics; see Heckman (1981) for an important early contribution along these lines.

to choose the sector again. This raises concerns because knowledge of steady-state worker flows is no longer sufficient to characterize the effect of sectoral shocks.[20]

In response to this issue, we provide suggestive evidence that points to the importance of worker heterogeneity in generating the gap in shown in Figure 3. Our strategy involves constructing an alternative proxy for worker heterogeneity. Instead of relying on demographic and socioeconomic characteristics, we leverage workers' sector choice histories prior to 1990 as a means to capture their heterogeneity. Since differences among workers manifest as differences in their sector choice patterns, these histories allow us to more effectively control for their heterogeneity. Specifically, we use pre-1990 data to compute each worker's one-year manufacturing staying probability and then categorize workers into five types based on the quintiles of that probability. Next, we assess the extent to which accounting for these worker types narrows the gap between actual and model-implied staying probabilities. In Figure 6, the actual k-year staying probabilities calculated using post-1990 data are denoted by blue dots, while red and green lines represent the staying probabilities implied by the canonical model and the model with the five worker types, respectively.

If the gap were driven entirely by duration dependence mechanism, the green line would coincide with the red line, because our construction only accounts for the effect of worker heterogeneity, while leaving the duration dependence mechanism unchanged. However, the green line is close to the blue dots, reducing the gap by more than half. Given that sector choice history only partially captures worker heterogeneity, this finding suggests that at least half of the gap is attributable to worker heterogeneity—the mechanism emphasized in this paper.

4.4 Extrapolation of Worker Flow Matrices

While we could only calculate a finite number of worker flow matrices, sufficient statistics approach requires a full sequence of worker flow matrices from k equals one to infinity. Thus, we need a method to extrapolate longer-run worker flow matrices from the available finite-length data.[21] Leveraging the structural model provides a natural method for extrapolation.[22] For this purpose, we estimate the structural model by matching the worker flow matrices we computed directly from the NLSY data, $\{\mathcal{F}_k\}_{k=1}^{18}$. See Section OA.2.1 for details. The estimated structural model generates the full set of worker flow matrices, which are used in Section 5 to perform counterfactual exercises. Since the structural model reduces to the canonical model when there is only one worker type, this extrapolation is a strict generalization of the canonical model's extrapolation.

Surprisingly, the model with only two worker types (i.e., $|\Omega| = 2$) closely matches the observed worker flow matrices. Figures 3 and 4 document the fit of the estimated model. The blue solid line in Figure 3 represents the model-implied k-year staying probabilities for the manufacturing sector, which is almost indistinguishable from the actual staying probabilities in the data (blue dots). In Figure 4, we use blue dots to

[20] Note that the sufficient statistics result on welfare changes, derived from the envelope theorem, extends to a much broader class of models, including those with duration dependence mechanism.

[21] With $\beta = 0.96$, what happens after $k = 18$ years is not negligible, as $0.96^{18} \approx 0.5$.

[22] This approach has the advantage that extrapolation is disciplined by the model. However, the sufficient statistics approach does not, in principle, require that we estimate the details of the model. In Section OA.2.2, we explore alternative extrapolation methods that do not rely on a structural model.

plot all 4×4 elements of the model-implied worker flow matrix series against the actual values in the data. Most of the dots lie roughly on the 45-degree line.

4.5 Estimation of ρ

In this section, we present a novel strategy for estimating the parameter ρ, which, in conjunction with the extrapolated worker flow matrix series, provides a complete description of the labor supply side of the model. Our goal is to propose an estimation method that does not require explicitly specifying the underlying heterogeneity. The possibility of such an estimation is already suggested in the second equation of Proposition 1, which we restate here for convenience:

$$\mathrm{d}\ln \ell_t = \sum_{s \geq 0, k \geq 0} \frac{\beta^{k+1}}{\rho} (\mathcal{F}_{s+k} - \mathcal{F}_{s+k+2}) \, \mathbb{E}_{t-s-1} \, \mathrm{d}w_{t-s+k}. \tag{15}$$

This equation describes the response of sectoral employment to wage shocks, with the coefficients depending only on the value of ρ and the worker flow matrix series, independent of the specific details of worker heterogeneity. Thus, we can estimate ρ by measuring the responsiveness of sectoral employment to sectoral wage shocks, conditional on the observed worker flow matrix series.

Implementation. In principle, equation (15) could be directly confronted with data to estimate ρ, but this requires that we fully specify how much information workers have about the future. The literature circumvents this demanding requirement by applying the Euler equation approach first used in ACM (e.g., Artuç and McLaren, 2015; Caliendo, Dvorkin, and Parro, 2019; Traiberman, 2019). We extend the Euler equation approach by allowing for arbitrary worker heterogeneity.

The key idea of the Euler equation approach is to transform (15) into a recursive representation. For example, in the absence of persistent worker heterogeneity, we can use the restriction $\mathcal{F}_k = (\mathcal{F}_1)^k$ to rewrite equation (15) recursively as follows:[23]

$$\mathrm{d}\ln \ell_t = (\mathcal{F}_1^{-1} + \beta \mathcal{F}_1) \, \mathrm{d}\ln \ell_{t+1} - \beta \, \mathbb{E}_t \, \mathrm{d}\ln \ell_{t+2} - \frac{\beta}{\rho} (\mathcal{F}_1^{-1} - \mathcal{F}_1) \, \mathbb{E}_t \, \mathrm{d}w_{t+1}. \tag{16}$$

See Section OA.2.4 for proofs of the results in this section. However, with arbitrary worker heterogeneity, this is impossible because the worker flow matrix series that determines the coefficients of equation (15) is based on empirical data and does not have a recursive structure. Nevertheless, we will demonstrate in two steps that it is still possible to derive an *approximate* recursive representation of equation (15). First, we show that when there is a finite number of worker types, N, equation (15) always possesses an *exact* recursive representation of the form

$$\mathrm{d}\ln \ell_t = \sum_{k=1}^{4N-2} \Gamma_k \, \mathbb{E}_t \, \mathrm{d}\ln \ell_{t+k} + \frac{\beta}{\rho} \sum_{k=1}^{4N-3} \Lambda_k \, \mathbb{E}_t \, \mathrm{d}w_{t+k}, \tag{17}$$

[23] Unlike equation (15), equation (16) contains only the term $\mathbb{E}_t \, \mathrm{d}w_{t+1}$ since the (expected) values of $\mathrm{d}\ln \ell_{t+1}$ and $\mathrm{d}\ln \ell_{t+2}$ summarize the effect of all other beliefs. Because of this advantage, the Euler equation approach has been widely used in the literature, although previous studies have used a version of the equations that involves migration probabilities rather than labor supply.

where Γ_k and Λ_k are functions of worker flow matrix series, $\{\mathcal{F}_k\}$. Second, recall that a model with two worker types provides a close approximation to the observed worker flow matrix series $\{\mathcal{F}_k\}$ in the data. Combining these two findings, we can conclude that equation (15), with $\{\mathcal{F}_k\}$ from the data, can be approximately represented in the recursive form (17) with $N = 2$. In Section OA.2.4, we provide formulas for computing $\{\Gamma_k\}_{k=1,\ldots,6}$ and $\{\Lambda_k\}_{k=1,\ldots,5}$. We use simulation to demonstrate that the obtained approximate recursive representation provides a close fit to the actual dynamics of sectoral employment; see Figure OA.8.

We further modify the obtained recursive representation in two ways:

$$\ln \ell_t - \sum_{k=1}^{6} \Gamma_k \ln \ell_{t+k} = \frac{\beta}{\rho} \sum_{k=1}^{5} \Lambda_k w_{t+k} + \mathsf{ExpectationError}_{t+1,t+6}\,. \tag{18}$$

First, instead of deviations from steady-state values, $\mathrm{d}\ln\ell$ and $\mathrm{d}w$, we use the actual values, $\ln\ell$ and w. This is possible because the recursive representation always satisfies $\sum_{k=1}^{6} \Gamma_k = I$ and $\sum_{k=1}^{5} \Lambda_k = O$, where I is the identity matrix and O is the zero matrix. Second, the expected values are substituted with the realized values plus an expectation error term that depends on the news revealed between time $t+1$ and $t+6$.

Equation (18) is our regression specification, where we regress the left-hand-side variable on the explanatory variable on the right-hand side, $\sum_{k=1}^{5} \Lambda_k \,\mathrm{d}w_{t+k}$, to estimate $\frac{\beta}{\rho}$. However, since both the explanatory variable and the expectation error term are affected by newly revealed information between period $t+1$ and $t+6$, they are likely to be correlated. To address this concern, we follow ACM and use past values of sectoral labor allocation and wages as instruments. Any variables included in the period t information set are theoretically valid instruments for the explanatory variable, providing a consistent estimate of ρ. In particular, we use the 1-year lag of sectoral wages and sectoral employment shares. For this approach to be valid, it is necessary to assume that workers have rational expectations and that the error term in equation (18) only reflects errors in workers' forecasts. If this term also incorporates shocks to the labor supply curve (e.g., unexpected aggregate shifts in preferences for particular sectors), we must assume an exclusion restriction whereby such shocks are uncorrelated with our instruments. For a discussion of the strengths and weaknesses of this approach in the case of homogeneous workers, refer to ACM and Traiberman (2019). Also, we include sector fixed effects in the regression to isolate within-sector variation from across-sector variation.

Our estimation method requires data on sectoral employment and sectoral wages. We use the Bureau of Labor Statistics' Current Employment Surveys (CES) to compute the time series of these variables. We use the share of workers in the dataset employed in each sector as our measure of sectoral labor allocation and average wages in each sector as our measure of sectoral wages. We again consider four broad sectors. In Section 5, we will compare the counterfactual predictions of our model with those of Caliendo, Dvorkin, and Parro (2019), who analyze a homogeneous counterpart of our model. To facilitate clear comparison, we assume that the variable w represents the logarithm of the sectoral wage.

Table 1 presents the estimation results. Column (1) estimates equation (18) by OLS. The estimated coefficient and the implied value of ρ are negative and insignificant. In Column (2), we estimate the same specification using IV. The implied value of ρ is 0.825. This means that a one standard deviation higher realization of the idiosyncratic shock is associated with 4.06% higher lifetime consumption (see Section OA.2.3). Comparing the results in Columns (1) and (2), it appears that the estimated coefficient $\widehat{\beta/\rho}$ from OLS

Table 1: Estimation of ρ

	(1)	(2)	(3)
β/ρ	-0.286	1.164^{**}	0.877^{***}
	(0.311)	(0.550)	(0.296)
Implied ρ	-3.358	0.825^{**}	1.095^{***}
	(3.652)	(0.390)	(0.369)
Method	OLS	IV	IV
Persistent Heterogeneity	✓	✓	
No Persistent Heterogeneity			✓
Observations	136	136	136
First-stage F	–	38.7	11.1

Notes: OLS and IV estimation results for specification (18) (heterogeneous workers: Columns (1) and (2)) and (16) (homogeneous workers: Column (3)). Standard errors robust to heteroskedasticity are reported in parentheses, with ** $p < 0.05$, *** $p < 0.01$. Data source: NLSY79 and BLS.

is biased downward due to the presence of expectation errors. This is consistent with the fact that expectation errors are likely to cause sectoral labor supply and sectoral wages to be negatively correlated.[24] The estimate in Column (2) implies that a temporary 1% decline in manufacturing wages leads to a 0.35% decrease in the manufacturing share, while a permanent 1% decline in manufacturing wages is associated with a 1.15% decrease in the manufacturing share. In column (3), we assume that workers are homogeneous and estimate equation (16) by IV with the same set of instrument variables. The implied value of ρ is slightly higher than our preferred estimate, but we cannot reject the null of equality between the two. Estimates in Columns (2) and (3) are broadly consistent with the estimates of ρ in the literature, which range from 0.5 to 2. The original ACM and subsequent papers estimate ρ to be around two. A relatively more recent paper, Artuç and McLaren (2015), suggests a value of $\rho = 0.62$. Also, Rodriguez-Clare, Ulate, and Vasquez (2022) obtain a similar value of $\rho = 0.56$.[25] This estimate in Column (2) is our preferred specification, and we will use it in the subsequent analysis.

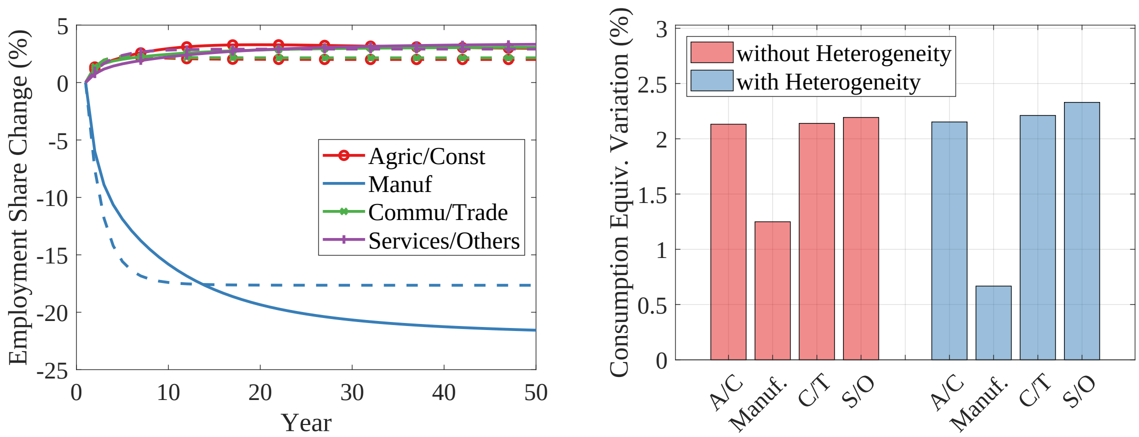
We now move on to applications of our model, where we use our empirical estimates to quantify the effect of sectoral shocks.

5. Applications

In this section, we consider two sectoral shocks and quantify the implications of persistent worker heterogeneity for welfare and labor reallocation. We first apply our results to a stylized trade liberalization exercise of

[24] If workers wrongly expect an increase in wages in a sector, they would supply more labor to that sector. This increased labor supply would lead to a decrease in wages in that sector.

[25] Note that ACM and other papers in the literature assume instantaneous utility linear in wage. This implies that the value of ρ governs semi-elasticity $\frac{\partial \ln \ell_{t+k}}{\partial \text{wage}_t}$ instead of elasticity $\frac{\partial \ln \ell_{t+k}}{\partial \ln \text{wage}_t}$. However, because they normalize sectoral wages so that the average annualized wage equals unity, both the semi-elasticity and elasticity can be interpreted as the percentage change in sectoral employment in response to a percentage change in sectoral wages.



(a) Changes in Sectoral Employment Share

(b) Changes in Sectoral Values

Figure 7. Counterfactual Changes in Sectoral Employment and Welfare: Trade Liberalization

Notes: This figure plots the transitional dynamics following an unexpected permanent drop in manufacturing prices. Solid lines and solid bars correspond to the prediction from the sufficient statistics in the data, and dashed lines and hatched bars correspond to the prediction of the canonical model without persistent worker heterogeneity.

ACM, which clearly illustrates the biases of the canonical model and the underlying mechanism. For a more realistic application, we then examine the dynamic effects of the rise in China's import competition on US labor markets. We revisit this extensively studied topic using the sufficient statistics approach. Although we focus on these two exercises, our result can be applied more generally to other papers in the literature.

5.1 Permanent Decline in Manufacturing Prices

The first counterfactual exercise closely follows ACM and considers an unanticipated permanent 10% drop in manufacturing prices for a small open economy—for example, due to trade liberalization. Following Section 2, we incorporate persistent worker heterogeneity in the labor supply side of the model, which is calibrated in Section 4. We specify the labor demand side of the model as in ACM. We assume log utility with Cobb-Douglas consumption aggregate and a sector-specific CES production function with fixed capital stock. All goods are traded and their prices are exogenously given at world price level. Sectoral wages are competitively determined by the marginal productivity of labor. The labor demand side of the model is calibrated exactly as in ACM; see Section OA.3.1 for details. Initially, the economy is in a steady state. Since wages are endogenously determined by the labor market equilibrium, we apply Proposition 2 to compute the perfect-foresight transition path following announcement of the shock in year 1 until the economy reaches a new steady state.

Figure 7 shows results of the counterfactual exercise. In Figure 7a, we plot the dynamics of sectoral employment predicted by the models with and without worker heterogeneity. The manufacturing employment share drops sharply, by around 20%. Importantly, the canonical model overestimates the short-term labor reallocation, resulting in a much faster transition to the new steady state. The transition is completed within 10 years in the absence of worker heterogeneity, while it takes more than 50 years with worker heterogeneity.

At the same time, the canonical model underestimates the magnitude of long-term reallocation. All of these results are consistent with the predictions of Proposition 4.[26]

In Figure 7b, we plot changes in welfare, measured in terms of consumption-equivalent variation; that is, the proportional change in the lifetime consumption sequence that would have the same effect on household welfare as would the welfare effect of the decline in manufacturing prices. In particular, the last four bars represent the welfare changes of workers based on their initial sector of employment predicted by the model with worker heterogeneity. The first four bars represent welfare changes implied by the canonical model. The result shows that even workers who were initially employed in manufacturing—the import-competing sector—benefit from trade liberalization. The increase in option value, driven by the increase in real wages in other sectors, more than compensates for the decline in manufacturing wages.[27] However, consistent with the prediction of Proposition 3, the canonical model overestimates the gains of manufacturing workers by a factor of almost two. At the same time, it slightly underestimates the gains of non-manufacturing sector workers, resulting in substantial underestimation of the distributional consequences of trade liberalization.

As discussed in Section 3, the welfare gains of manufacturing workers are overestimated in the canonical model for two reasons. Not only does it overestimate welfare gains for given wage changes, but it also predicts a smaller decline in manufacturing wages in the short term. To illustrate this, we show that if we used the changes in sectoral wages computed from the canonical model—instead of endogenizing them—and combined them with the worker flow matrix series as in Proposition 1, we would get a narrower gap between the welfare changes predicted by models with and without worker heterogeneity; see Figure OA.10.

To gain deeper understanding of the disparities between models with and without heterogeneity, we plot changes in sectoral values and employment shares separately for each of the two worker types in Figure OA.11. When the shock hits the economy, type 1 workers, who have lower switching costs and comparative advantage in non-manufacturing sectors, can easily move out of manufacturing and enjoy a higher welfare gain from the shock. Over time, as more type 2 workers leave the manufacturing sector, manufacturing wages begin to recover. Given that type 2 workers are more likely to be stuck in the manufacturing sector once they choose it, they dislike manufacturing more than type 1 workers, resulting in a higher proportion of type 1 workers in the manufacturing sector in the long run.

Finally, we assess the quality of the first-order approximation around a steady state. In Figure OA.12, we compare the results of the counterfactual exercises obtained using sufficient statistics formulas with those calculated from the exact solution of the estimated structural model.[28] Despite the relatively large magnitude of the 10% shock considered in this section, the sufficient statistics results deliver a close approximation. The

[26] In Section OA.3.1, we compute the impulse responses of sectoral employment from the observed worker flow matrix series. As predicted by Proposition OA.2, the canonical model tends to overestimate the short-term impact of shocks on sectoral employment but underestimates their long-term effects. Notably, within a 7-year time horizon, the canonical model consistently overestimates the effects of shocks on sectoral employment. This explains the difference between changes in sectoral employment in the models with and without worker heterogeneity documented in Figure 7a.

[27] The real wage of the manufacturing sector initially overshoots because it takes time for labor to adjust. As the number of workers in manufacturing gradually declines, manufacturing wages rise and eventually exceed the preshock steady-state level. However, they increase less than those in other sectors. Thus, the manufacturing employment share declines over time and in the new steady state. See Figure OA.9 for details.

[28] The figure also demonstrates that the exact value changes always exceed those calculated using a first-order approximation. We can show analytically that the second-order term is always positive due to the option value.

approximation error becomes almost negligible for a shock size of 1%, but becomes more pronounced as the shock size increases to 30%.

5.2 The China Shock

As a second application, we apply our sufficient statistics result to a more realistic counterfactual exercise: the dynamic impact of the growth of China's manufacturing productivity and the resulting import competition on the welfare of US workers and labor reallocation. We closely follow the workhorese dynamic quantitative trade model of Caliendo, Dvorkin, and Parro (2019) (hereafter, CDP) and extend the labor supply side by allowing for arbitrary time-invariant worker heterogeneity.[29] We make two simplifying assumptions relative to the original specification of CDP. First, we consider four broad sectors—Manufacturing, Wholesale/Retail, Construction, and Services—and another auxiliary sector representing nonemployment.[30] Second, for reasons discussed shortly, we abstract from interstate migration and assume that workers can switch sectors only within each state.[31,32] We first describe the labor supply side, then specify the labor demand side and the shock of interest to close the model.

Labor Supply Side. Two observations motivate us to conduct a separate analysis for each of the 50 US states. First, it is well known that there is considerable variation in exposure to the China shock across different geographic regions in the US (see, for example, Autor, Dorn, and Hanson, 2013; Acemoglu et al., 2016). This suggests that Assumption 1 is better imposed within each state. Second, worker flow matrices differ significantly across states; workers in states with a higher manufacturing employment share are more likely to remain in the manufacturing sector over time. By applying the sufficient statistics approach at state level, we can account for such state-level heterogeneity. Under our simplifying assumptions, we can focus on 5-by-5 intersectoral worker flow matrices for each state.

State-level analysis requires computation of a worker flow matrix series for each state. However, the limited sample size of the NLSY data makes it difficult to estimate them accurately. Thus, we instead use the monthly Current Population Survey (CPS) dataset, which contains a substantial sample size for each state. This dataset tracks workers for 4 consecutive months, which allows us to compute worker flow matrices

[29] A large body of subsequent literature studies additional elements that are missing in this framework: involuntary unemployment from downward nominal wage rigidities or search frictions (Kim and Vogel, 2021; Rodriguez-Clare, Ulate, and Vasquez, 2022); endogenous trade imbalances (Dix-Carneiro et al., 2023); occupation adjustment (Traiberman, 2019); agglomeration forces and employment sensitivity to wages (Adao, Arkolakis, and Esposito, 2023); learning and expectations (Fan, Hong, and Parro, 2023); and currency pegs (Kim, de la Barrera, and Fukui, 2023). Incorporating worker heterogeneity in models with these additional features is an important direction for future research.

[30] In CDP, there are 23 sectors: 12 from Manufacturing, 8 from Services, and 1 each for Wholesale/Retail, Construction, and nonemployment.

[31] CDP uses an 1150-by-1150 quarterly worker flow matrix between all US state-sector pairs (50 states, excluding the District of Columbia and the territories, and 22 sectors plus 1 additional sector representing non-employment). However, we need to estimate the longer-run worker flow matrix as well as the short-run one, and estimating them at this level of granularity is practically impossible.

[32] In principle, allowing for regional migration would dampen the welfare effects because it provides an additional margin of adjustment. However, US workers change sectors almost 10 to 100 times more often than they change states, which suggests that the majority of the labor adjustment occurs at the sector change margin. In the same context of the China shock, Rodriguez-Clare, Ulate, and Vasquez (2022) also show that ignoring migration makes little difference for their model's prediction. Autor, Dorn, and Hanson (2013) and Autor et al. (2014) also find little evidence that regional migration is an important mechanism through which the economy adjusts to the China shock.

$\mathcal{F}_k^{\text{state}}$ for $k = 1, 2, 3$ months for each state. Specifically, we assume that the US was in a steady state before the China shock and compute worker flows by pooling transition observations between January 1995 and December 1999.[33] To compute worker flow matrices for other values of k, we follow Section 4 and estimate the structural model with two types of workers by matching the observed worker flow matrices. We then use the estimated model to extrapolate longer-run worker flow matrix series.[34] In Figure OA.16, we compare the fits of the models with and without worker heterogeneity to the observed worker flow matrix series. While the fits of the models vary across states, the canonical model consistently underestimates the staying probabilities. Finally, we use the value of the parameter ρ estimated in Section 4.[35]

Labor Demand Side and the China Shock. The labor demand side of the model is more complex than in Section 5.1: It features a large number of labor markets distinguished by sector and region, international trade, interregional trade within the US, input-output linkages, and multiple production inputs. Section OA.3.2 describes the primitive assumptions of the model regarding households' consumption and sector choices; intermediate goods and final goods producing firms' profit maximization; and the definition of a sequential competitive equilibrium. We follow CDP in calibrating the structural parameters of the labor demand side; see Section OA.3.2 for details of the calibration. The shock of interest is the growth of China's manufacturing productivity. Following CDP, we consider the China shock to be a sequence of shocks to the growth rate of total factor productivity (TFP) of the Chinese manufacturing sector from 2000 to 2007, assuming a constant fundamental thereafter. We also assume that US agents anticipated the China shock in 2000 exactly as it occurred. We calibrate manufacturing productivity growth such that the model's predicted increase in US imports from China exactly matches the predicted increase in imports, using the increase in imports from China of the other eight advanced economies as an instrument. See Section OA.3.1 for detailed calibration of the China shock.

Counterfactual Results. For exogenous changes in the manufacturing productivity of China, US sectoral wages are endogenously determined by the labor market equilibrium. Thus, we again use Proposition 2 to calculate counterfactual changes in sectoral welfare and sectoral employment. Given the rich structure of the model, it is computationally demanding to estimate all exogenous state variables of the model—including productivities, labor mobility costs, and trade costs—for every period. We reduce the computational burden by extending the CDP's dynamic hat algebra to models with arbitrary worker heterogeneity using our sufficient

[33] The monthly CPS dataset also tracks workers again for 4 additional consecutive months after 8 months after the first 4 consecutive months. Thus, in principle, we can observe $\mathcal{F}_k^{\text{state}}$ for $k = 1, 2, 3, 12, 13, 14, 15$. However, it is well known that the monthly CPS dataset suffers from underestimation of staying probabilities when comparing the first 4 months with the second 4 months because sectors are coded independently between these months (e.g., Kambourov and Manovskii, 2013). This can be clearly seen in Figure OA.14, where we plot the manufacturing staying probabilities. This underestimation is critical to our analysis, so we focus only on the first three worker flow matrices. This also minimizes concerns about sample attrition and the resulting selection (e.g., Moscarini and Thomsson, 2007).

[34] One concern is that this requires too much extrapolation. In Figure OA.15, we compare the extrapolated state-specific worker flow matrix series with the aggregate worker flow matrix series we computed in Section 4. Reassuringly, the average state-specific worker flow matrix series behaves very similarly to the aggregate series.

[35] In Section 4.5, we estimate the parameter ρ at annual frequency. In Section OA.3.2, we present a way to transform this to quarterly frequency. The resulting value is $\rho = 1.0011$.

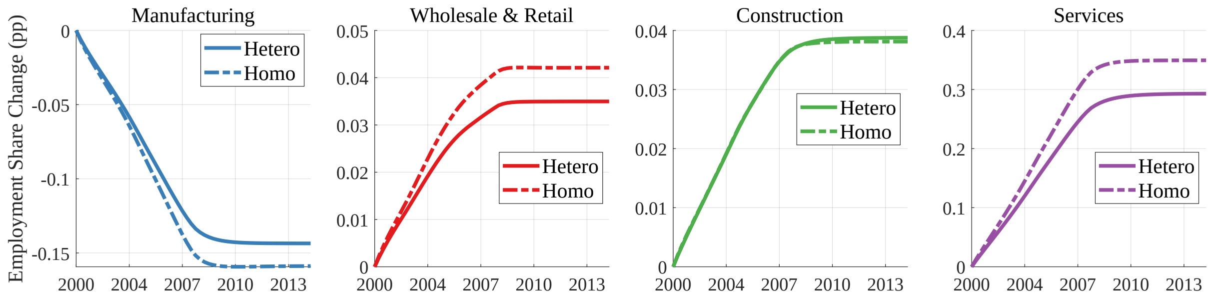


Figure 8. Effect of the China Shock on Sectoral Employment Share

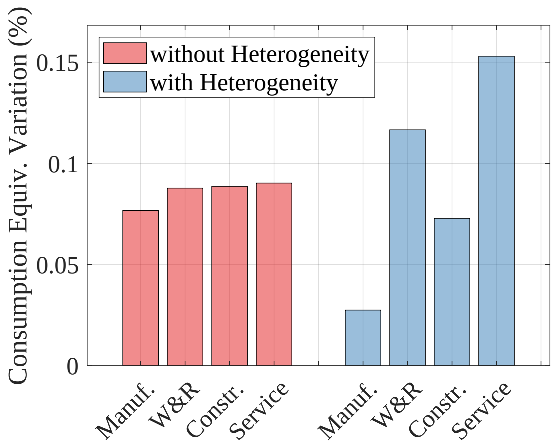
Notes: This figure plots the transitional dynamics following the China shock. Solid lines correspond to the prediction from the sufficient statistics in the data, and dashed lines correspond to the prediction of the canonical model without persistent worker heterogeneity.

statistics result; see Section OA.3.2 for details.[36] Figures 8 and 9 plot the results of counterfactual exercises computed using the sufficient statistics approach.

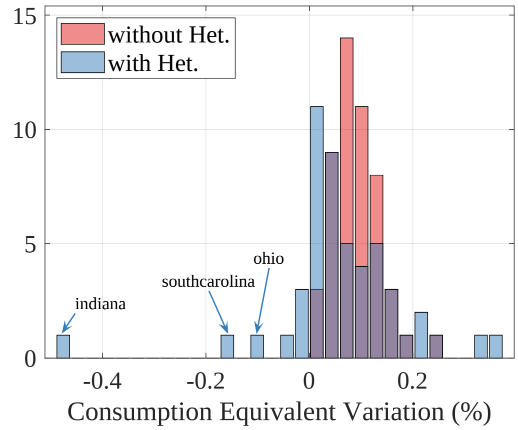
Sectoral Employment Changes. Figure 8 plots the dynamic response of US sectoral employment to the China shock for models with and without worker heterogeneity. In both models, the increase in the manufacturing productivity of China shifts US employment from manufacturing to non-manufacturing sectors, which benefit from access to cheaper intermediate goods made available by the China shock. With worker heterogeneity, the China shock reduces the manufacturing employment share by around 0.14 percentage points (equivalent to 0.32 million jobs) over the 10-year period, while increasing the employment share of services by around 0.3 percentage points. Dashed lines plot changes in the sectoral employment share predicted by the canonical model. As expected from Proposition 4, the canonical model consistently overestimates the extent of labor reallocation by up to 20%.

Welfare Changes. In terms of aggregate welfare, the heterogeneous-worker model predicts a 0.086% increase (in terms of consumption-equivalent variation), which is similar to the 0.101% welfare change predicted by the canonical model. Despite this similarity, the two models differ in their predictions regarding the distributional conflicts—how much the winners gain and the losers lose. In Figure 9a, we plot sectoral welfare changes measured at the time the China shock was known to the US labor market. Again, the last four bars represent sectoral welfare changes predicted using the sufficient statistics in the data, and the first four bars represent those implied by the canonical model. As in Section 5.1, both models predict that even workers who were initially employed in the manufacturing sector benefit from the China shock. However, as expected from the results of Proposition 3, the canonical model significantly overestimates the welfare gain of manufacturing workers, but underestimates the welfare gain of workers in non-manufacturing sectors. Thus, the model significantly underestimates the distributional impact of the China shock. In particular,

[36] Dynamic hat algebra solves the equilibrium of the model in terms of time differences and differences between the actual and counterfactual economies. This method allows one to perform counterfactual exercises without the need to estimate the level of exogenous state variables of the model. For this reason, it is widely used in the literature (e.g., Rodriguez-Clare, Ulate, and Vasquez, 2022; Caliendo et al., 2021; Balboni, 2019; Kleinman, Liu, and Redding, 2023).



(a) Sectoral Welfare Changes: Aggregate Effect



(b) Manufacturing Welfare Changes: Individual States

Figure 9. Effect of the China Shock on Sectoral Employment

Notes: This figure plots the sectoral welfare changes following the China shock. In the left panel, solid bars correspond to the prediction from the sufficient statistics in the data, and hatched bars correspond to the prediction of the canonical model without persistent worker heterogeneity. In the right panel, we plot a histogram of state-level changes in the welfare of workers initially employed in manufacturing.

in the absence of worker heterogeneity, workers who were initially employed in the manufacturing sector enjoy about the same welfare gains as non-manufacturing workers. However, once we account for worker heterogeneity and correctly match longer-run worker flow patterns, their average gain becomes less than a quarter of the gains of workers in the other sectors.

The welfare impact of China's import competition varies substantially across regions. In Figure 9b, we present a histogram of state-level changes in the welfare of manufacturing workers. In the canonical model, manufacturing workers in all states benefit from the China shock. In contrast, the heterogeneous-worker model predicts that the welfare gain of manufacturing workers is close to zero in most states, and manufacturing workers from seven states—Alabama, Florida, Indiana, Ohio, Rhode Island, South Carolina, and Vermont—experience a decline in welfare. These states have much higher manufacturing employment shares than other states and experience higher reallocation from manufacturing to nonemployment. The figure also shows that worker heterogeneity not only amplifies the negative welfare effects of the China shock but also increases regional disparities in the welfare effects on manufacturing workers.

We also plot the percentage changes in (real) sectoral wages averaged across states induced by the China shock in Figure OA.17. The average effect is negative for the manufacturing sector. In contrast, non-manufacturing sectors experience wage increases in all states. More importantly, the heterogeneous-worker model predicts slower labor reallocation from manufacturing to non-manufacturing, resulting in a larger decline in manufacturing wages. Motivated by this observation, we plot sectoral welfare changes calculated by combining the worker flow matrix series implied by the heterogeneous-worker model and sectoral real wage changes implied by the model without worker heterogeneity in Figure OA.18. More than half of the welfare change gap between the two models arises from this difference in real wage changes. This highlights the importance of endogenizing wage changes when studying the role of worker heterogeneity.

6. Concluding Remarks

Large economic shocks often have asymmetric effects across different sectors. Such shocks can necessitate substantial labor reallocation across sectors and may have significant distributional consequences for workers employed in different sectors. The key determinant of the dynamic effects of sectoral shocks is the ease with which workers can switch sectors over time. In this paper, we develop a dynamic sector choice model that incorporates a self-selection mechanism based on persistent worker heterogeneity in an otherwise standard dynamic discrete choice framework. Our sufficient statistics approach, which relies on the information contained in steady-state worker flows over various horizons, highlights in a transparent way the critical role of this self-selection mechanism in shaping the dynamic effects of sectoral shocks. Assuming away persistent worker heterogeneity results in overestimation of steady-state worker flows, which in turn leads to underestimation of the distributional consequences and overestimation of the speed of labor reallocation.

By revisiting the two applications in the literature using our empirical estimates of the sufficient statistics, we illustrate the applicability of our approach and quantify the importance of the added flexibility from worker heterogeneity. Our results present a more pessimistic view of the consequences of sectoral shocks: The reallocation of workers is significantly slower, and the welfare losses of adversely affected workers are more severe than previously suggested. Although we have focused on specific applications in this paper, the general insights we have developed could be applied more broadly—not only to other sectoral shocks but also to models that incorporate richer mechanisms. For example, if involuntary unemployment is considered, our finding that workers are more likely to be stuck in a negatively affected sector will materialize as a higher unemployment rate in that sector, leading to even greater welfare losses. Applying our approach in the context of different shocks and to models with additional structures is an important direction for future research.

Acknowledgements. Jaeeun Seo is deeply indebted to Arnaud Costinot, George-Marios Angeletos, and Dave Donaldson for their invaluable guidance and support. We also thank Daron Acemoglu, Fernando Alvarez, Costas Arkolakis, David Atkin, Martin Beraja, Ricardo Caballero, Bumsoo Kim, Michael Peters, Ivan Werning, Christian Wolf, and participants at Berkeley, LSE, MIT, Northwestern, 2024 SED conference, Tokyo, and Yale for their helpful comments. All errors are our own.

References

Acemoglu, Daron, David Autor, David Dorn, Gordon H Hanson, and Brendan Price (2016) "Import Competition and the Great US Employment Sag of the 2000s," *Journal of Labor Economics*, 34 (S1), S141–S198.

Acemoglu, Daron, and Pascual Restrepo (2020) "Robots and Jobs: Evidence From US Labor Markets," *Journal of political economy*, 128 (6), 2188–2244.

Adão, Rodrigo (2016) "Worker Heterogeneity, Wage Inequality, and International Trade: Theory and Evidence from Brazil," *Working Paper* (November).

Adao, Rodrigo, Costas Arkolakis, and Federico Esposito (2023) "General Equilibrium Effects in Space: Theory and Measurement."

Adao, Rodrigo, Martin Beraja, and Nitya Pandalai-Nayar (2023) "Fast and Slow Technological Transitions,"Technical report, Tech. rep., MIT, Mimeo.

Aguirregabiria, Victor, and Pedro Mira (2010) "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 156 (1), 38–67.

Allen, Treb, and Costas Arkolakis (2014) "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129 (3), 1085–1140.

Allen, Treb, Costas Arkolakis, and Yuta Takahashi (2020) "Universal Gravity," *Journal of Political Economy*, 128 (2), 393–433.

Allen, Treb, and Dave Donaldson (2020) "Persistence and Path Dependence in the Spatial Economy,"Technical report, National Bureau of Economic Research.

Amior, Michael, and Alan Manning (2018) "The Persistence of Local Joblessness," *American Economic Review*, 108 (7), 1942–1970.

Arcidiacono, Peter, and Robert A Miller (2011) "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 79 (6), 1823–1867.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare (2012) "New trade models, same old gains?," *American Economic Review*, 102 (1), 94–130.

Artuç, Erhan, Shubham Chaudhuri, and John McLaren (2010) "Trade shocks and labor adjustment: A structural empirical approach," *American Economic Review*, 100 (3), 1008–1045.

Artuç, Erhan, and John McLaren (2015) "Trade policy and wage inequality: A structural analysis with occupational and sectoral mobility," *Journal of International Economics*, 97 (2), 278–294.

Autor, David H, David Dorn, and Gordon H Hanson (2013) "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American economic review*, 103 (6), 2121–2168.

Autor, David H, David Dorn, Gordon H Hanson, and Jae Song (2014) "Trade Adjustment: Worker-level Evidence," *The Quarterly Journal of Economics*, 129 (4), 1799–1860.

Balboni, Clare Alexandra (2019) *In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities* Ph.D. dissertation, London School of Economics and Political Science.

Baqaee, David Rezza, and Emmanuel Farhi (2020) "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135 (1), 105–163.

Beraja, Martin (2023) "A Semistructural Methodology for Policy Counterfactuals," *Journal of Political Economy*, 131 (1), 190–201.

Bilal, Adrien, and Esteban Rossi-Hansberg (2021) "Location As an Asset," *Econometrica*, 89 (5), 2459–2495.

Borjas, George J (1987) "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 531–553.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel (2019) "Changes in Between-group Inequality: Computers, Occupations, and International Trade," *American Economic Journal: Macroeconomics*, 11 (2), 348–400.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro (2019) "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87 (3), 741–835.

Caliendo, Lorenzo, Luca David Opromolla, Fernando Parro, and Alessandro Sforza (2021) "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement," *Journal of Political Economy*, 129 (12), 3491–3545.

Chetty, Raj (2009) "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-form Methods," *Annu. Rev. Econ.*, 1 (1), 451–488.

Chetty, Raj, John N Friedman, Michael Stepner et al. (2020) "The Economic Impacts of COVID-19: Evidence From a New Public Database Built Using Private Sector Data,"Technical report, national Bureau of economic research.

Costinot, Arnaud, and Jonathan Vogel (2010) "Matching and Inequality in the World Economy," *Journal of Political Economy*, 118 (4), 747–786.

Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg (2018) "The Geography of Development," *Journal of Political Economy*, 126 (3), 903–983.

Dix-Carneiro, Rafael (2014) "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 82 (3), 825–885.

Dix-Carneiro, Rafael, João Paulo Pessoa, Ricardo Reyes-Heroles, and Sharon Traiberman (2023) "Globalization, Trade Imbalances, and Labor Market Adjustment," *The Quarterly Journal of Economics*, 138 (2), 1109–1171.

Dvorkin, Maximiliano A (2023) "International trade and labor reallocation: misclassification errors, mobility, and switching costs."

Eaton, Jonathan, and Samuel Kortum (2002) "Technology, Geography, and Trade," *Econometrica*, 70 (5), 1741–1779.

Fan, Jingting, Sungwan Hong, and Fernando Parro (2023) "Learning and Expectations in Dynamic Spatial Economies,"Technical report, National Bureau of Economic Research.

French, Eric, and Christopher Taber (2011) "Identification of Models of the Labor Market," in *Handbook of labor economics*, 4, 537–617: Elsevier.

Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi (2023) "Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade," *The Review of Economic Studies*, 90 (1), 331–375.

Goldberg, Pinelopi Koujianou, and Nina Pavcnik (2007) "Distributional Effects of Globalization in Developing Countries," *Journal of economic Literature*, 45 (1), 39–82.

Grigsby, John R (2022) “Skill Heterogeneity and Aggregate Labor Market Dynamics,”Technical report, National Bureau of Economic Research.

Heckman, James J (1981) “Heterogeneity and State Dependence,” in *Studies in labor markets*, 91–140: University of Chicago Press.

Heckman, James J, and Bo E Honore (1990) “The Empirical Content of the Roy Model,” *Econometrica: Journal of the Econometric Society*, 1121–1149.

Heckman, James J, and Guilherme Sedlacek (1985) “Heterogeneity, Aggregation, and Market Wage Functions: an Empirical Model of Self-selection in the Labor Market,” *Journal of political Economy*, 93 (6), 1077–1125.

Heckman, James, and Burton Singer (1984) “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data,” *Econometrica: Journal of the Econometric Society*, 271–320.

Howard, Greg, and Hansen Shao (2023) “The Dynamics of Internal Migration: A New Fact and Its Implications.”

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow (2019) “The Allocation of Talent and Us Economic Growth,” *Econometrica*, 87 (5), 1439–1474.

Hulten, Charles R (1978) “Growth Accounting with Intermediate Inputs,” *The Review of Economic Studies*, 45 (3), 511–518.

Kambourov, Gueorgui, and Iourii Manovskii (2013) “A Cautionary Note On Using (March) Current Population Survey and Panel Study of Income Dynamics Data To Study Worker Mobility,” *Macroeconomic Dynamics*, 17 (1), 172–194.

Keane, Michael P, and Eswar S Prasad (1996) “The Employment and Wage Effects of Oil Price Changes: A Sectoral Analysis,” *The Review of Economics and Statistics*, 389–400.

Keane, Michael P, and Kenneth I Wolpin (1997) “The Career Decisions of Young Men,” *Journal of political Economy*, 105 (3), 473–522.

Kennan, John, and James R Walker (2011) “The Effect of Expected Income On Individual Migration Decisions,” *Econometrica*, 79 (1), 211–251.

Kim, Bumsoo, Marc de la Barrera, and Masao Fukui (2023) “Currency Pegs, Trade Imbalances and Unemployment: A Reevaluation of the China Shock.”

Kim, Ryan, and Jonathan Vogel (2020) “Trade and Welfare (Across Local Labor Markets),” (April).

——— (2021) “Trade Shocks and Labor Market Adjustment,” *American Economic Review: Insights*, 3 (1), 115–130.

Kleinman, Benny, Ernest Liu, and Stephen J Redding (2023) “Dynamic Spatial General Equilibrium,” *Econometrica*, 91 (2), 385–424.

Lagakos, David, and Michael E Waugh (2013) “Selection, Agriculture, and Cross-country Productivity Differences,” *American Economic Review*, 103 (2), 948–980.

Lee, Donghoon (2005) “An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice,” *International Economic Review*, 46 (1), 1–34.

Lee, Donghoon, and Kenneth I Wolpin (2006) “Intersectoral Labor Mobility and the Growth of the Service Sector,” *Econometrica*, 74 (1), 1–46.

Lee, Eunhee (2020) "Trade, Inequality, and the Endogenous Sorting Ofheterogeneous Workers," *Journal of International Economics*, 125, 103310.

Lise, Jeremy, and Fabien Postel-Vinay (2020) "Multidimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*.

McFadden, Daniel (1973) "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontier in Econometrics*.

McKay, Alisdair, and Christian K Wolf (2022) "What Can Time-Series Regressions Tell Us About Policy Counterfactuals?,"Technical report, National Bureau of Economic Research.

Milgrom, Paul, and Ilya Segal (2002) "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70 (2), 583–601.

Moscarini, Giuseppe, and Kaj Thomsson (2007) "Occupational and Job Mobility in the US," *The Scandinavian Journal of Economics*, 109 (4), 807–836.

Pierce, Justin R, and Peter K Schott (2016) "The Surprisingly Swift Decline of US Manufacturing Employment," *American Economic Review*, 106 (7), 1632–1662.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo (2022) "The Human Side of Structural Transformation," *American Economic Review*, 112 (8), 2774–2814.

Rodriguez-Clare, Andres, Mauricio Ulate, and Jose P Vasquez (2022) "Trade with Nominal Rigidities: Understanding the Unemployment and Welfare Effects of the China Shock."

Rust, John (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica: Journal of the Econometric Society*, 999–1033.

Sprung-Keyser, Ben, Nathaniel Hendren, and Sonya Porter (2022) *The Radius of Economic Opportunity: Evidence From Migration and Local Labor Markets*: US Census Bureau, Center for Economic Studies.

Traiberman, Sharon (2019) "Occupations and import competition: Evidence from Denmark," *American Economic Review*, 109 (12), 4260–4301.

Young, Alwyn (2014) "Structural Transformation, the Mismeasurement of Productivity Growth, and the Cost Disease of Services," *American Economic Review*, 104 (11), 3635–67.

Appendix

A. Theoretical Results

A.1 Aggregation Result: Proposition 1

In this section, we first prove a general result that does not rely on Assumption 2. In the next section, we show how Assumption 2 simplifies it to Proposition 1. To state the general result, we define another type of worker flow matrices. As we will demonstrate, two types of worker flow matrices are sufficient statistics for characterizing the welfare and labor market consequences of sectoral shocks.

Definition A.1. *For each $m, k \in \mathbb{N}$, the (m, k)-period* worker flow matrix*, is an $S \times S$ matrix denoted as $\mathcal{F}_{m,k}$ whose (i, j)-element is given by*

$$(\mathcal{F}_{m,k})_{i,j} = \Pr(s_{\tau(t,m)+k} = j | s_t = i),$$

where the random variable $\tau(t, m) \equiv \min\{\tau \geq t : s_\tau = s_{t-m}\}$ denotes the first period in which a worker returns to the sector she chose m periods ago.

The (i, j)-element of the (m, k)-period worker flow matrix equals the steady-state probability that a randomly selected worker from sector i will move to sector j after k periods after returning to the sector she chose m periods ago.

To formalize the idea of aggregation, we define the population-average operator.

Definition A.2. *The* population-average operator $\bar{\mathbb{E}}_\omega$ *is an operator that can be applied to $S \times N$ matrices for any $N \in \mathbb{N}$. It maps a type-specific matrix to an aggregated sector-level matrix,*

$$\begin{pmatrix} a^\omega_{11} & \cdots & a^\omega_{1N} \\ \vdots & \ddots & \vdots \\ a^\omega_{S1} & \cdots & a^\omega_{SN} \end{pmatrix} \mapsto \int_\Omega \begin{pmatrix} a^\omega_{11}\, \mathrm{d}W(\omega|s=1) & \cdots & a^\omega_{1N}\, \mathrm{d}W(\omega|s=1) \\ \vdots & \ddots & \vdots \\ a^\omega_{S1}\, \mathrm{d}W(\omega|s=S) & \cdots & a^\omega_{SN}\, \mathrm{d}W(\omega|s=S) \end{pmatrix},$$

where $\mathrm{d}W(\omega|s=i) = \frac{\ell^\omega_i\, \mathrm{d}W(\omega)}{\int_\Omega \ell^{\omega'}_i\, \mathrm{d}W(\omega')}$ represents the steady-state type distribution of sector i.

If the i-th row of a matrix contains variables related to sector i, then the steady-state type distribution of sector i gives the appropriate weights for computing the average across different types. This is precisely how the population-average operator is defined. The following lemma shows that certain type-specific variables can be converted to their aggregate equivalents through application of the population-average operator.[37] We relegate the proof into Appendix B along all other proofs omitted in the appendix.

[37] The population-average operator does not transform all type-specific variables into their aggregate counterparts. For example, nonlinear functions of $\mathrm{d}\ln \ell^\omega_{t+1}$, $\mathrm{d}v^\omega_t$, or $(F^\omega)^k$ do not possess this property (e.g., $\bar{\mathbb{E}}_\omega\, \mathrm{d}\ell^\omega_{t+1} \neq \mathrm{d}\ell_{t+1}$).

Lemma A.1. *If we apply the population-average operator to* $\mathrm{d}\ln\ell^{\omega}_{t+1}$, $\mathrm{d}v^{\omega}_{t}$, $(F^{\omega})^{k}$, *or* $(B^{\omega})^{m}(F^{\omega})^{k}$, *we obtain aggregated sector-level variables:*

$$\bar{\mathbb{E}}_{\omega}\,\mathrm{d}\ln\ell^{\omega}_{t+1}=\mathrm{d}\ln\ell_{t+1},\quad \bar{\mathbb{E}}_{\omega}\,\mathrm{d}v^{\omega}_{t}=\mathrm{d}v_{t},\quad \bar{\mathbb{E}}_{\omega}\big[(F^{\omega})^{k}\big]=\mathcal{F}_{k},\quad \bar{\mathbb{E}}_{\omega}\big[(B^{\omega})^{m}(F^{\omega})^{k}\big]=\mathcal{F}_{m,k}.$$

In particular, with infinite-length longitudinal information on workers' sector choices, we can observe $\bar{\mathbb{E}}_{\omega}[(F^{\omega})^{k}]$ and $\bar{\mathbb{E}}_{\omega}[(B^{\omega})^{m}(F^{\omega})^{k}]$ for all $k,m\in\mathbb{N}_0$.[38] We can now apply the population-average operator to the left- and right-hand sides of equations (7) and (8) to derive a general version of the sufficient statistics result. First, we have

$$\mathrm{d}v_t=\bar{\mathbb{E}}_{\omega}\,\mathrm{d}v^{\omega}_{t}=\bar{\mathbb{E}}_{\omega}\Big[\sum_{k\geq 0}(\beta F^{\omega})^{k}\,\mathbb{E}_t\,\mathrm{d}w_{t+k}\Big]=\sum_{k\geq 0}\beta^{k}\bar{\mathbb{E}}_{\omega}\Big[(F^{\omega})^{k}\Big]\,\mathbb{E}_t\,\mathrm{d}w_{t+k}=\sum_{k\geq 0}\beta^{k}\mathcal{F}_k\,\mathbb{E}_t\,\mathrm{d}w_{t+k}.$$

Second, we have

$$\begin{aligned}\mathrm{d}\ln\ell_t=\bar{\mathbb{E}}_{\omega}\,\mathrm{d}\ln\ell^{\omega}_{t}&=\bar{\mathbb{E}}_{\omega}\left[\frac{\beta}{\rho}\sum_{s\geq 0}(B^{\omega})^{s}(I-B^{\omega}F^{\omega})\left(\sum_{k\geq 0}(\beta F^{\omega})^{k}\,\mathbb{E}_{t-s-1}\,\mathrm{d}w^{\omega}_{t-s+k}\right)\right]\\&=\bar{\mathbb{E}}_{\omega}\left[\sum_{s\geq 0,k\geq 0}\frac{\beta^{k+1}}{\rho}\big((B^{\omega})^{s}(F^{\omega})^{k}-(B^{\omega})^{s+1}(F^{\omega})^{k+1}\big)\,\mathbb{E}_{t-s-1}\,\mathrm{d}w_{t-s+k}\right]\\&=\sum_{s\geq 0,k\geq 0}\frac{\beta^{k+1}}{\rho}(\mathcal{F}_{s,k}-\mathcal{F}_{s+1,k+1})\,\mathbb{E}_{t-s-1}\,\mathrm{d}w_{t-s+k}.\end{aligned}$$

The result is summarized in the following proposition.

Proposition A.1 (Sufficient Statistics Result without Assumption 2). *Suppose that Assumption 1 holds. For a given sequence of (common) changes in sectoral wages* $\{\mathrm{d}w_t\}$*, the changes in sectoral value and sectoral employment are given by*

$$\begin{aligned}\mathrm{d}v_t&=\sum_{k\geq 0}\beta^{k}\mathcal{F}_k\,\mathbb{E}_t\,\mathrm{d}w_{t+k},\\ \mathrm{d}\ln\ell_t&=\sum_{s\geq 0,k\geq 0}\frac{\beta^{k+1}}{\rho}(\mathcal{F}_{s,k}-\mathcal{F}_{s+1,k+1})\,\mathbb{E}_{t-s-1}\,\mathrm{d}w_{t-s+k}.\end{aligned}$$

Proposition A.1 establishes that in order to construct the counterfactual changes in welfare and sectoral employment for a given sequence of sectoral wage changes, we only require knowledge of the two types of worker flow matrices, $\{\mathcal{F}_k\}$ and $\{\mathcal{F}_{s,k}\}$, along with the parameters β and ρ. Proposition 2 can be generalized in a similar manner.

[38] For the second type of worker flow matrices, this requires an additional assumption that $(\mathcal{F}_k)_{i,j}$ is strictly positive for all $i,j\in\mathcal{S}$, and $k\in\mathbb{N}$.

A.2 Simplified Sufficient Statistics Result under Assumption 2

Lemma A.1 demonstrates that with infinite-length longitudinal information on workers' sector choices, we can observe two types of worker flow matrices, $\{\mathcal{F}_k\}$ and $\{\mathcal{F}_{m,k}\}$, for all k and m. In practice, however, we can only observe them for small enough k and m due to the finite nature of the real-world datasets. In this section, we show how these data requirements can be significantly reduced under Assumption 2.

Lemma A.2. *Under Assumption 2, we have $B^\omega = F^\omega$ for all $\omega \in \Omega$. In this case, we have $\mathcal{F}_{m,k} = \mathcal{F}_{m+k}$.*

Assumption 2 imposes a certain structure on the sector-switching costs. Lemma A.2 shows that this assumption implies the equality between the backward transition matrix and the forward transition matrix, which in turn implies that the second type of worker flow matrices reduces to the first type of worker flow matrices. Thus, we can focus on one type of worker flow matrices, and Proposition A.1 is simplified to Proposition 1 in the main text.

In the left panel of Figure OA.6, we compute the transition matrices between the four sectors considered in the main text and plot the 16 elements of the backward transition matrices against the corresponding elements of the forward transition matrices. All elements lie closely on the 45-degree line. It is not possible to test this implication empirically at the unobserved type level. Instead, we compute the backward and forward transition matrices for observed types. Specifically, we consider four dimensions of observed heterogeneity, as in Section 4: gender, race, education, and age, which results in 16 groups. In the right panel of Figure OA.6, we compare the matrices for all 16 groups. Once again, we observe a high degree of similarity between the backward and forward transition matrices.

A.3 First-Order Approximation of Labor Demand Side

Suppose that type-specific sectoral wages are determined by the labor allocation and exogenous shocks:

$$w_{it}^{\omega} = f_i^{\omega}\left(\{\ell_{jt}\}_j, \{\xi_{jt}\}_j\right)$$

for $i \in \mathcal{S}$ and $\omega \in \Omega$. Up to a first-order approximation, we can write

$$\mathrm{d}w_{it}^{\omega} = \sum_j \frac{\partial f_i^{\omega}}{\partial \ln \ell_{jt}} \cdot \mathrm{d}\ln \ell_{jt} + \sum_j \frac{\partial f_i^{\omega}}{\partial \xi_{jt}} \mathrm{d}\xi_{jt}.$$

Assumption 1 requires that $\mathrm{d}w_{it}^{\omega} = \mathrm{d}w_{it}$ for all $\omega \in \Omega$ for any realizations of $\{\ell_{jt}\}_j$ and $\{\xi_{jt}\}_j$. This in turn requires that

$$\mathrm{d}w_{it}^{\omega} = \mathrm{d}w_{it} \equiv \sum_j \frac{\partial f_i^{\omega_1}}{\partial \ln \ell_{jt}} \cdot \mathrm{d}\ln \ell_{jt} + \sum_j \frac{\partial f_i^{\omega_1}}{\partial \xi_{jt}} \mathrm{d}\xi_{jt}$$

for all $\omega \in \Omega$ for a given $\omega_1 \in \Omega$.[39] Thus, we can write

$$\begin{aligned} \mathrm{d}w_t &= \begin{pmatrix} \frac{\partial f_1^{\omega_1}}{\partial \ln \ell_{1t}} & \frac{\partial f_1^{\omega_1}}{\partial \ln \ell_{2t}} & \cdots & \frac{\partial f_1^{\omega_1}}{\partial \ln \ell_{St}} \\ \frac{\partial f_2^{\omega_1}}{\partial \ln \ell_{1t}} & \frac{\partial f_2^{\omega_1}}{\partial \ln \ell_{2t}} & \cdots & \frac{\partial f_2^{\omega_1}}{\partial \ln \ell_{St}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_S^{\omega_1}}{\partial \ln \ell_{1t}} & \frac{\partial f_S^{\omega_1}}{\partial \ln \ell_{2t}} & \cdots & \frac{\partial f_S^{\omega_1}}{\partial \ln \ell_{St}} \end{pmatrix} \mathrm{d}\ln \ell_t + \begin{pmatrix} \frac{\partial f_1^{\omega_1}}{\partial \xi_{1t}} & \frac{\partial f_1^{\omega_1}}{\partial \xi_{2t}} & \cdots & \frac{\partial f_1^{\omega_1}}{\partial \xi_{St}} \\ \frac{\partial f_2^{\omega_1}}{\partial \xi_{1t}} & \frac{\partial f_2^{\omega_1}}{\partial \xi_{2t}} & \cdots & \frac{\partial f_2^{\omega_1}}{\partial \xi_{St}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_S^{\omega_1}}{\partial \xi_{1t}} & \frac{\partial f_S^{\omega_1}}{\partial \xi_{2t}} & \cdots & \frac{\partial f_S^{\omega_1}}{\partial \xi_{St}} \end{pmatrix} \mathrm{d}\varepsilon_t. \\ &= D \cdot \mathrm{d}\ln \ell_t + E \cdot \mathrm{d}\xi_t. \end{aligned}$$

A.4 Intuition for the Sufficient Statistics Result

We begin by providing intuition at the micro (i.e., type) level, then show how this intuition is preserved at the macro level when type-specific equations are aggregated across different types of workers.

Intuition at the Micro Level. Key equations at the micro level are (5) and (6). Equation (5) is an application of the envelope theorem (e.g., Milgrom and Segal, 2002). The envelope theorem implies that changes in future optimal sector choices do not contribute to the change in current welfare. Thus, we can evaluate the effect of changes in future sectoral wages using workers' transition probabilities as weights. For example, equation (7) shows that the effect of a unit change in the period-$(t+k)$ sector j wage on the period-t welfare of sector i workers is given by the probability that these workers transition to sector j in period $t+k$ (after discounting the future):

$$\frac{\partial v_{it}^{\omega}}{\partial \mathbb{E}_t w_{j,t+k}^{\omega}} = \beta^k \Pr(s_{t+k} = j | s_t = i, \omega) \equiv \beta^k \left[(F^{\omega})^k\right]_{ij}.$$

Thus, the matrix $(F^{\omega})^k$ is—by providing information about workers' k-period sector choices—informative about the effect of sectoral shocks on welfare.

Equation (6) summarizes the labor reallocation in response to sectoral shocks, the lagged version of which is rewritten here for convenience:

$$\mathrm{d}\ln \ell_t^{\omega} = B^{\omega}\, \mathrm{d}\ln \ell_{t-1}^{\omega} + \tfrac{\beta}{\rho^{\omega}} (I - B^{\omega} F^{\omega})\, \mathbb{E}_{t-1}\, \mathrm{d}v_t^{\omega}. \tag{6}$$

The first term captures the mechanical effect of changes in labor allocation in the previous period, and the second term captures the response of workers' sector choices to changes in sectoral values in period t. In particular, it implies

$$\left.\frac{\partial \ln \ell_{it}^{\omega}}{\partial \mathbb{E}_t v_{jt}^{\omega}}\right|_{\text{fix } \ell_{t-1}} = \tfrac{\beta}{\rho} \Big(\mathbb{1}_{i=j} - \left(B^{\omega} F^{\omega}\right)_{ij} \Big).$$

[39] Note that this does not necessarily imply that $w_{it}^{\omega} = w_{it}$ for all $\omega \in \Omega$.

To understand this equation, suppose first that $i \neq j$. The semi-elasticity of employment in sector i with respect to the expected value of sector j is proportional to the (i, j) element of $B^{\omega} F^{\omega}$. Note that we have

$$(B^{\omega} F^{\omega})_{ij} = \sum_{k \in \mathcal{S}} B^{\omega}_{ik} F^{\omega}_{kj} = \sum_{k \in \mathcal{S}} \Pr(s_{t-1} = k | s_t = i, \omega) \Pr(s_t = j | s_{t-1} = k, \omega).$$

Thus, this element can be interpreted as the probability that a type ω worker who chooses $s_t = i$ will switch to $s_t = j$ if they are allowed to choose period-t sector again after redrawing period-t idiosyncratic shocks. If this probability is high for a given sector pair (i, j), it means that many workers are at the margin between sectors i and j. In this case, a small change in the value of sector j leads to a relatively larger decline in employment in sector i.

For $i = j$, an increase in the value of sector i leads to a relatively larger increase in sector i employment when many workers are at the margin between sector i and other sectors, which is measured by $1 - (B^{\omega} F^{\omega})_{ii}$.

Moving to the longer run, we assume $\beta = 1$ to make the intuition clear, while the intuition remains the same for $\beta < 1$. Consider a change in v^{ω}_{jt} that is known to workers in period $t - k$, $k \in \mathbb{N}$. The effect of this news on ℓ^{ω}_{it} can be written as follows:

$$\begin{aligned} \frac{\partial \ln \ell^{\omega}_{it}}{\partial \, \mathbb{E}_{t-k} \, v^{\omega}_{jt}} &= \tfrac{1}{\rho} \Big((I - B^{\omega} F^{\omega}) + (B^{\omega} F^{\omega} - (B^{\omega})^2 (F^{\omega})^2) + \cdots + ((B^{\omega})^{k-1} (F^{\omega})^{k-1} - (B^{\omega})^k (F^{\omega})^k) \Big)_{ij} \\ &= \tfrac{1}{\rho} \Big(\mathbb{1}_{i=j} - \big((B^{\omega})^k (F^{\omega})^k \big)_{ij} \Big). \end{aligned}$$

Again, the (i, j)-element of the matrix $(B^{\omega})^k (F^{\omega})^k$ can be interpreted as the probability that workers who initially choose sector i in period t will switch to sector j, if they are allowed to choose sectors again for periods $t + 1 - k, t + 2 - k, \cdots, t$ after redrawing idiosyncratic shocks for these periods. This provides a measure of how many workers are at the margin between sectors i and j over the k-period horizon, thereby being informative about how responsive workers are to shocks known k periods in advance.

In this sense, the steady-state transition matrices and their kth powers are at least qualitatively informative about the response of sectoral employment to sectoral shocks. However, the sufficient statistics results rely on the semi-elasticity being an exact (linear) function of products of the transition matrices B^{ω} and F^{ω}—a condition guaranteed under the extreme-value distribution assumption.

Intuition at the Macro Level. At the micro level, the intuition can be expressed in terms of (products of) transition matrices. Appendix A.1 demonstrates that we can aggregate type-specific equations to derive macro-level equations, establishing the sufficient statistics result. In particular, Lemma A.2 shows that the products of transition matrices can be aggregated to two types of worker flow matrices. This preserves the same intuition at the macro level: The matrix $\mathcal{F}_k$ is—by providing information about the *average* worker's k-period sector choices—informative about the effect of sectoral shocks on welfare; and the matrix $\mathcal{F}_{s,k}$, which is equal to $\mathcal{F}_{s+k}$ under Assumption 2, is—by providing information about how many workers are at the margin between each pair of sectors—informative about the response of sectoral employment to sectoral shocks.

B. Omitted Proofs

Proof of Equations (5) and (6). Taking the first-order approximations to equations (2) and (4) and the definition of F_{ijt}, we have

$$
\begin{aligned}
\mathrm{d}v_{it} &= \mathrm{d}w_{it} + \rho \sum_j \left. \frac{\Big(\exp(\beta\,\mathbb{E}_t\, v_{jt+1})/\exp(C_{ij})\Big)^{1/\rho}}{\sum_k \Big(\exp(\beta\,\mathbb{E}_t\, v_{kt+1})/\exp(C_{ik})\Big)^{1/\rho}} \right|_{\text{steady state}} \cdot \tfrac{\beta}{\rho}\,\mathbb{E}_t\,\mathrm{d}v_{jt+1} \\
&= \mathrm{d}w_{it} + \beta \sum_j F_{ij}\,\mathbb{E}_t\,\mathrm{d}v_{jt+1}, \qquad \text{(A.1)}\\
\mathrm{d}\ln \ell_{jt+1} &= \sum_i \left. \frac{F_{ijt}\ell_{it}}{\ell_{jt+1}} \right|_{\text{steady state}} \cdot (\mathrm{d}\ln \ell_{it} + \mathrm{d}\ln F_{ijt}) = \sum_i B_{ji} \cdot (\mathrm{d}\ln \ell_{it} + \mathrm{d}\ln F_{ijt}), \qquad \text{(A.2)}\\
\mathrm{d}\ln F_{ijt} &= \tfrac{\beta}{\rho}\left(\mathbb{E}_t\,\mathrm{d}v_{jt+1} - \sum_k F_{ik}\,\mathbb{E}_t\,\mathrm{d}v_{kt+1}\right). \qquad \text{(A.3)}
\end{aligned}
$$

Plugging equation (A.3) into equation (A.2), we have

$$
\begin{aligned}
\mathrm{d}\ln \ell_{jt+1} &= \sum_i B_{ji}\,\mathrm{d}\ln \ell_{it} + \tfrac{\beta}{\rho} \sum_i B_{ji} \cdot \left(\mathbb{E}_t\,\mathrm{d}v_{jt+1} - \sum_k F_{ik}\,\mathbb{E}_t\,\mathrm{d}v_{kt+1}\right) \\
&= \sum_i B_{ji}\,\mathrm{d}\ln \ell_{it} + \tfrac{\beta}{\rho} \cdot \left(\mathbb{E}_t\,\mathrm{d}v_{jt+1} - \sum_i B_{ji} \sum_k F_{ik}\,\mathbb{E}_t\,\mathrm{d}v_{kt+1}\right). \qquad \text{(A.4)}
\end{aligned}
$$

With vector notation, equations (A.1) and (A.4) become equations (5) and (6). □

Proof of Lemma 1. Equation (5) relates $\mathrm{d}v_t^\omega$ with $\mathrm{d}v_{t+1}^\omega$. Solving this equation forward, we can write $\mathrm{d}v_t^\omega$ as a linear combination of the expected value of a sequence $(\mathrm{d}w_t^\omega, \mathrm{d}w_{t+1}^\omega, \mathrm{d}w_{t+2}^\omega, \cdots)$,

$$
\mathrm{d}v_t^\omega = \sum_{k\geq 0} (\beta F^\omega)^k\,\mathbb{E}_t\,\mathrm{d}w_{t+k}^\omega.
$$

Plugging this result into equation (6), we obtain a formula that relates $\mathrm{d}\ln \ell_{t+1}^\omega$ with $\mathrm{d}\ln \ell_t^\omega$,

$$
\mathrm{d}\ln \ell_{t+1}^\omega = B^\omega\,\mathrm{d}\ln \ell_t^\omega + \tfrac{\beta}{\rho}(I - B^\omega F^\omega) \sum_{k\geq 0} (\beta F^\omega)^k\,\mathbb{E}_t\,\mathrm{d}w_{t+k+1}^\omega.
$$

Solving this equation backward, we can write $\mathrm{d}\ln \ell_t^\omega$ as a linear combination of the expected value of a two-sided sequence $(\cdots, \mathrm{d}w_{t-1}^\omega, \mathrm{d}w_t^\omega, \mathrm{d}w_{t+1}^\omega, \cdots)$,

$$
\mathrm{d}\ln \ell_t^\omega = \tfrac{\beta}{\rho} \sum_{s\geq 0} (B^\omega)^s (I - B^\omega F^\omega) \left(\sum_{k\geq 0} (\beta F^\omega)^k\,\mathbb{E}_{t-s-1}\,\mathrm{d}w_{t-s+k}^\omega\right).
$$

□

Proof of Lemma A.2. Evaluating equation (3) at steady state, we have

$$F_{ij}^{\omega} = \frac{\left(\exp(\beta v_j^{\omega})/\exp(C_{ij}^{\omega})\right)^{1/\rho^{\omega}}}{\sum_{k\in\mathcal{S}}(\exp(\beta v_k^{\omega})/\exp(C_{ik}^{\omega}))^{1/\rho^{\omega}}}.$$

For notational convenience, write $\theta_j = \exp(\frac{\beta}{\rho^{\omega}} v_j^{\omega})$ and $c_{ij} = \exp(\frac{1}{\rho^{\omega}} C_{ij}^{\omega})$, then $F_{ij}^{\omega} = \frac{\theta_j/c_{ij}}{\sum_k \theta_k/c_{ik}}$. Under Assumption 2, we can write $c_{ij} = c_i \cdot \tilde{c}_j \cdot \hat{c}_{ij}$ where $\hat{c}_{ij} = \hat{c}_{ji}$ for all $i, j \in \mathcal{S}$. Thus,

$$F_{ij}^{\omega} = \frac{\frac{\theta_j}{c_i\tilde{c}_j\hat{c}_{ij}}}{\sum_k \frac{\theta_k}{c_i\tilde{c}_k\hat{c}_{ik}}} = \frac{\frac{\theta_j/\tilde{c}_j}{\hat{c}_{ij}}}{\sum_k \frac{\theta_k/\tilde{c}_k}{\hat{c}_{ik}}} = \frac{\tilde{\theta}_j/\hat{c}_{ij}}{\sum_k \tilde{\theta}_k/\hat{c}_{ik}} = \frac{\tilde{\theta}_j/\hat{c}_{ij}}{\varphi_i},$$

where $\tilde{\theta}_i \equiv \theta_i/\tilde{c}_i$ and $\varphi_i \equiv \sum_k \tilde{\theta}_k/\hat{c}_{ik}$. We can easily check that $\ell_i^{\omega} = \frac{\tilde{\theta}_i\varphi_i}{\sum_k \tilde{\theta}_k\varphi_k}$ solves the equation $\ell_i^{\omega} = \sum_k \ell_k^{\omega} F_{ki}^{\omega}$, $\forall i$. Thus, we have balanced bilateral worker flows at steady state,

$$\frac{\ell_i^{\omega} F_{ij}^{\omega}}{\ell_j^{\omega} F_{ji}^{\omega}} = \frac{\tilde{\theta}_i\varphi_i \frac{\tilde{\theta}_j/\hat{c}_{ij}}{\varphi_i}}{\tilde{\theta}_j\varphi_j \frac{\tilde{\theta}_i/\hat{c}_{ji}}{\varphi_j}} = 1.$$

Thus, we have $B_{ji}^{\omega} = \frac{\ell_i^{\omega} F_{ij}^{\omega}}{\ell_j^{\omega}} = F_{ji}^{\omega}$ for all $i, j \in \mathcal{S}$. □

Proof of Lemma 2. When there are two sectors, Section OA.1.3 shows

$$(\mathcal{F}_k)_{11} = 1 - \frac{\sum_{\omega} \tilde{\beta}_{\omega}\theta_{\omega}\alpha_{\omega} f^k(2-\alpha_{\omega}-\beta_{\omega})}{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}},$$

$$\left((\mathcal{F}_1)^k\right)_{11} = 1 - \frac{\sum_{\omega} \tilde{\beta}_{\omega}\theta_{\omega}\alpha_{\omega}}{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}} f^k\left(2 - \frac{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}\alpha_{\omega}}{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}} - \frac{\sum_{\omega}\tilde{\alpha}_{\omega}\theta_{\omega}\beta_{\omega}}{\sum_{\omega}\tilde{\alpha}_{\omega}\theta_{\omega}}\right),$$

where $F^{\omega} = \begin{pmatrix} 1-\alpha_{\omega} & \alpha_{\omega} \\ \beta_{\omega} & 1-\beta_{\omega} \end{pmatrix}$, $\tilde{\beta}_{\omega} = 1 - \tilde{\alpha}_{\omega} = \frac{\beta_{\omega}}{\alpha_{\omega}+\beta_{\omega}}$, $f^k(x) = \frac{1-(x-1)^k}{2-x}$, and θ_{ω} is the population share of type ω.

Thus, we need to show that

$$\frac{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}\alpha_{\omega}\tilde{f}^k(\alpha_{\omega}+\beta_{\omega})}{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}} < \frac{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}\alpha_{\omega}}{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}} \tilde{f}^k\left(\frac{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}\alpha_{\omega}}{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}} + \frac{\sum_{\omega}\tilde{\alpha}_{\omega}\theta_{\omega}\beta_{\omega}}{\sum_{\omega}\tilde{\alpha}_{\omega}\theta_{\omega}}\right),$$

where $\tilde{f}^k(x) = f^k(2-x)$, or equivalently

$$\frac{\sum_{\omega}\frac{\alpha_{\omega}\beta_{\omega}}{\alpha_{\omega}+\beta_{\omega}}\theta_{\omega}\tilde{f}^k(\alpha_{\omega}+\beta_{\omega})}{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}} < \frac{\sum_{\omega}\frac{\alpha_{\omega}\beta_{\omega}}{\alpha_{\omega}+\beta_{\omega}}\theta_{\omega}}{\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}} \tilde{f}^k\left(\frac{\sum_{\omega}\frac{\alpha_{\omega}\beta_{\omega}}{\alpha_{\omega}+\beta_{\omega}}\theta_{\omega}}{\left(\sum_{\omega}\tilde{\beta}_{\omega}\theta_{\omega}\right)\cdot\left(\sum_{\omega}\tilde{\alpha}_{\omega}\theta_{\omega}\right)}\right).$$

Define $g^k(x) = x \cdot \tilde{f}^k(x)$, this becomes

$$\frac{\sum_\omega \tilde{\alpha}_w \tilde{\beta}_w \theta_w g^k(\alpha_w + \beta_w)}{(\sum_\omega \tilde{\beta}_\omega \theta_\omega) \cdot (\sum_\omega \tilde{\alpha}_\omega \theta_\omega)} < g^k\left(\frac{\sum_\omega \tilde{\alpha}_w \tilde{\beta}_w \theta_\omega (\alpha_w + \beta_w)}{(\sum_\omega \tilde{\beta}_\omega \theta_\omega) \cdot (\sum_\omega \tilde{\alpha}_\omega \theta_\omega)}\right).$$

Define, $\tau_\omega = \frac{\tilde{\alpha}_\omega \tilde{\beta}_\omega \theta_\omega}{(\sum_\omega \tilde{\beta}_\omega \theta_\omega) \cdot (\sum_\omega \tilde{\alpha}_\omega \theta_\omega)}$, then $\kappa \equiv \sum_\omega \tau_w \in (0, 1)$ due to Jensen's inequality for $y \mapsto y^2$. Note that $g^k(x) = 1 - (1-x)^k$ is weakly increasing and weakly concave for $x \in [0, 1]$. Thus, we can apply Jensen's inequality two more times to have

$$\begin{aligned}
g^k\left(\frac{\sum_\omega \tilde{\alpha}_w \tilde{\beta}_w \theta_\omega (\alpha_w + \beta_w)}{(\sum_\omega \tilde{\beta}_\omega \theta_\omega) \cdot (\sum_\omega \tilde{\alpha}_\omega \theta_\omega)}\right) &= g^k\left(\sum_\omega \tau_w(\alpha_\omega + \beta_\omega)\right) = g^k\left(\kappa \sum_\omega \frac{\tau_\omega}{\kappa}(\alpha_w + \beta_w)\right) \\
&\geq \kappa g^k\left(\sum_\omega \frac{\tau_\omega}{\kappa}(\alpha_w + \beta_w)\right) > \kappa \sum_\omega \frac{\tau_\omega}{\kappa} g^k(\alpha_w + \beta_w) \\
&= \frac{\sum_\omega \tilde{\alpha}_w \tilde{\beta}_w \theta_w g^k(\alpha_w + \beta_w)}{(\sum_\omega \tilde{\beta}_\omega \theta_\omega) \cdot (\sum_\omega \tilde{\alpha}_\omega \theta_\omega)}.
\end{aligned}$$

□

Proof of Proposition 3. We can write the absolute changes in the value as

$$|\,\mathrm{d}v_{s1}| = \sum_{k \geq 0} \beta^k (\mathcal{F}_k)_{s,s} \Delta \geq \sum_{k \geq 0} \beta^k \left((\mathcal{F}_1)^k\right)_{s,s} \Delta = \left|\,\mathrm{d}v_{s1}\big|_{\text{canonical}}\right|.$$

□

Proof of Proposition 4. For $t \geq 2$,

$$\begin{aligned}
\mathrm{d}\ln \ell_t &= \sum_{s \geq 0, k \geq 0} \frac{\beta^{k+1}}{\rho}(\mathcal{F}_{s+k} - \mathcal{F}_{s+k+2})\, \mathbb{E}_{t-s-1}\, \mathrm{d}w_{t-s+k} \\
&= \sum_{s=0}^{t-2} \sum_{k \geq 0} \frac{\beta^{k+1}}{\rho}(\mathcal{F}_{s+k} - \mathcal{F}_{s+k+2})\, \mathrm{d}w_{t-s+k} \\
&= \sum_{k \geq 0} \frac{\beta^{k+1}}{\rho}(\mathcal{F}_k + \mathcal{F}_{k+1} - \mathcal{F}_{t+k-1} - \mathcal{F}_{t+k})\, \mathrm{d}w,
\end{aligned}$$

where $\mathrm{d}w = (0, \cdots, 0, \overbrace{-\Delta}^{s\text{th}}, 0, \cdots, 0)$. Thus, we can write

$$|\,\mathrm{d}\ln \ell_{st}| = \Delta \cdot \sum_{k \geq 0} \frac{\beta^{k+1}}{\rho}\left(\sum_{s=0}^{t-2} b_{s+k}\right).$$

This gives

$$|\,\mathrm{d}\ln \ell_{st+1}| - |\,\mathrm{d}\ln \ell_{st}| = \Delta \cdot \sum_{k \geq 0} \frac{\beta^{k+1}}{\rho} b_{k+t-1}.$$

Under the single-crossing condition of Assumption 3, if $|\,\mathrm{d}\ln \ell_{st+1}| - |\,\mathrm{d}\ln \ell_{st}|$ is higher in the heterogeneous-worker model, then $|\,\mathrm{d}\ln \ell_{st+2}| - |\,\mathrm{d}\ln \ell_{st+1}|$ is also higher in the heterogeneous-worker model. This means that there exists a cutoff $\bar{t} \in \mathbb{N} \cup \{\infty\}$ such that the canonical model calibrated by matching the one-period worker flow matrix overestimates the decline in employment of sector s in period t if and only if $1 < t \leq \bar{t}$. Note that

$$b_0 + \cdots + b_{t-2} = (I + \mathcal{F}_1 - \mathcal{F}_{t-1} - \mathcal{F}_t)_{ss}$$

is always higher in the canonical model. Thus, $|\,\mathrm{d}\ln \ell_{st}|$ is always higher in the canonical model if β is sufficiently close to zero. In that case, $\bar{t} = \infty$ is possible. On the other hand, for $t = 2$, we have

$$\begin{aligned}
|\,\mathrm{d}\ln \ell_{s2}| &= \Delta \cdot \sum_{k\geq 0} \frac{\beta^{k+1}}{\rho} (\mathcal{F}_k - \mathcal{F}_{k+2})_{s,s} \\
&= \Delta \cdot \frac{\beta}{\rho} \big((I - \mathcal{F}_2) + \beta(\mathcal{F}_1 - \mathcal{F}_3) + \beta^2(\mathcal{F}_2 - \mathcal{F}_4) + \cdots \big)_{s,s} \\
&= \Delta \cdot \frac{\beta}{\rho} \big(I + \beta\mathcal{F}_1 - (1-\beta^2)\mathcal{F}_2 - \beta(1-\beta^2)\mathcal{F}_3 - \cdots \big)_{s,s},
\end{aligned}$$

which is always higher in the canonical model. Thus, $\bar{t} \neq 1$. $\square$

Proof of Lemma A.1. The second equation holds by construction. For the first equation, we have

$$\begin{aligned}
[\bar{\mathbb{E}}_\omega \,\mathrm{d}\ln \ell^\omega_{t+1}]_i &= \int_\Omega \mathrm{d}\ln \ell^\omega_{it+1} \,\mathrm{d}W(\omega|s=i) \\
&= \frac{\int_\Omega \mathrm{d}\ln \ell^\omega_{it+1} \ell^\omega_i \,\mathrm{d}W(\omega)}{\ell_i} = \frac{\int_\Omega \mathrm{d}\ell^\omega_{it+1} \,\mathrm{d}W(\omega)}{\ell_i} = \frac{\mathrm{d}\ell_{it+1}}{\ell_i} = \mathrm{d}\ln \ell_{it+1}.
\end{aligned}$$

For the remaining equations, fix time t, we then have

$$\begin{aligned}
\big[\bar{\mathbb{E}}_\omega[(F^\omega)^k]\big]_{ij} &= \sum_\omega \tilde{\ell}^\omega_i \big[(F^\omega)^k\big]_{ij} = \sum_\omega \Pr(\omega|s_t = i) \cdot \Pr(s_t = i, s_{t+k} = j | s_t = i, \omega) \\
&= \sum_\omega \Pr(s_t = i, s_{t+k} = j, \omega | s_t = i) = \Pr(s_t = i, s_{t+k} = j | s_t = i) = (\mathcal{F}_k)_{ij}.
\end{aligned}$$

Again fix time t, and define a random variable $\tau(t,m)$ as follows:

$$\tau(t,m) \equiv \min\{\tau > t : s_\tau = s_{t-m}\},$$

which is well-defined based on two assumptions we begin with.

$$\begin{aligned}
\big[\bar{\mathbb{E}}_\omega[(B^\omega)^m (F^\omega)^k]\big]_{ij} &= \sum_\omega \tilde{\ell}^\omega_i \sum_h [(B^\omega)^m]_{ih} \cdot \big[(F^\omega)^k\big]_{hj} \\
&= \sum_\omega \Pr(\omega|s_t = i) \sum_h \Pr(s_{t-m} = h, s_t = i | s_t = i, \omega) \cdot \Pr(s_{\tau(t,m)+k} = j | s_{t-m} = h, s_t = i, \omega) \\
&= \sum_\omega \sum_h \Pr(s_{t-m} = h, s_t = i, s_{\tau(t,m)+k} = j, \omega | s_t = i) \\
&= \sum_\omega \Pr(s_t = i, s_{\tau(t,m)+k} = j, \omega | s_t = i) \\
&= \Pr(s_t = i, s_{\tau(t,m)+k} = j | s_t = i)
\end{aligned}$$

$\square$

C. Additional Figures

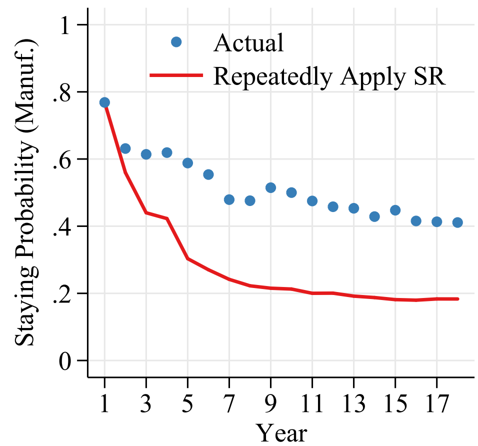


Figure A.1. Actual and Model-Implied Staying Probabilities: Non-Stationarity

Notes: For each $k = 1, \ldots, 18$ (year), this figure plots the actual probability of workers who choose manufacturing in 1980 choosing manufacturing again in year 1980+k (blue dots) and the probability computed by repeatedly multiplying time-varying one-year worker flow matrices (red line); i.e., the diagonal element of the matrix $\prod_{\kappa=1}^{k-1} \mathcal{F}_1^{\kappa}$ corresponding to the manufacturing sector. $\mathcal{F}_1^{\kappa}$ is the aggregate worker flow matrix computed using transition observations between years $1980 + \kappa - 1$ and $1980 + \kappa$ from the NLSY data.

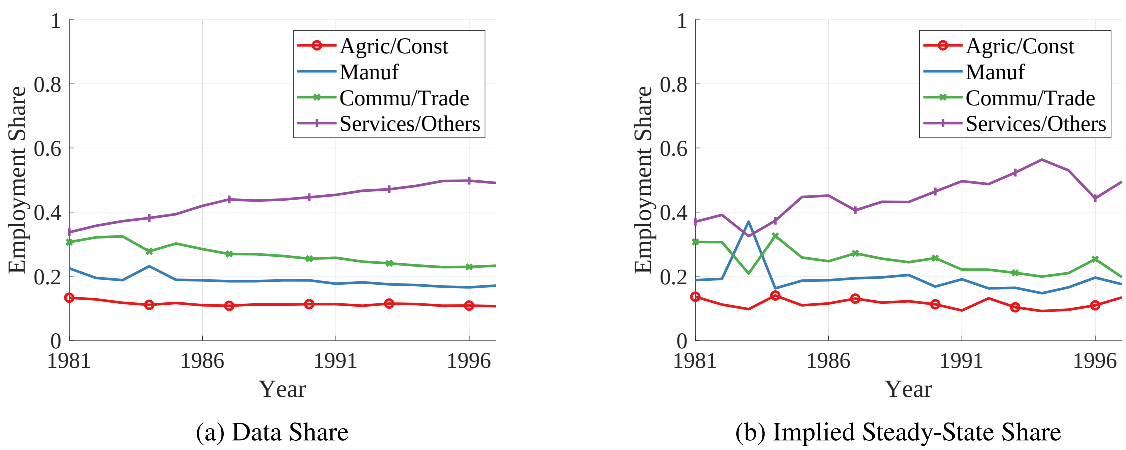


Figure A.2. Sectoral Employment Share

Notes: For each year between 1981 and 1997, this figure plots the sectoral employment share, both in the data (panel a) and the steady-state shares implied by the one-year worker flow matrix of that year (panel b). Data source: NLSY79.