Sectoral Shocks and Labor Market Dynamics: A Sufficient Statistics Approach*

Jaeeun Seo[†] Ryungha Oh[‡]

Job Market Paper

November 2023

Click here for the latest version

Abstract

In this paper, we introduce 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 accommodate arbitrary persistent heterogeneity across workers, we show that data on steady-state worker flows across sectors over different time horizons are sufficient to assess the labor supply responses to shocks as well as their welfare consequences, up to a first-order approximation. We also establish analytically that assuming away persistent worker heterogeneity, a common practice in existing literature, necessarily leads to overestimation of steady-state worker flows, resulting in systematic biases in counterfactual predictions. As an illustration of our sufficient statistics approach, we revisit the consequences of the rise of import competition from China. Using US panel data to measure steady-state worker flows prior to the shock, we find that the labor reallocation away from manufacturing is much slower, and the negative welfare effects on manufacturing workers are much more severe than those predicted by earlier models that abstract from persistent worker heterogeneity.

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

*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 in the MIT Macro Lunch and Trade lunch for their helpful comments. All errors are our own.

[†]Department of Economics, MIT, Cambridge, 02139, USA, *Email:* jaeeuns@mit.edu

[‡]Department of Economics, Yale University, New Haven, 06530, USA, *Email:* ryungha.oh@yale.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 very asymmetric across sectors, potentially creating both winners and losers. How much do the winners gain and the 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 in the long run, leading to more severe negative welfare consequences. Intuitively, these differences should be revealed by panel data on worker flows across sectors, i.e., workers' propensity to move over different time horizons. If so, one might be able to use such data as sufficient statistics for evaluating the impact of sectoral shocks on labor market dynamics and welfare. This paper formalizes this general intuition and demonstrates, both theoretically and empirically, the importance of persistent comparative advantage.

We focus on a broad class of dynamic discrete choice models that allow for arbitrary persistent heterogeneity across workers. At each point in time, workers decide in which sector to work and are subject to costs of switching sectors and transient idiosyncratic shocks drawn 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 productivity (or, more generally, in the utility they derive from being in a particular sector) and in the sector-switching costs. We impose no restriction on these persistent 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 provide a novel sufficient statistics approach that yields valid counterfactual predictions under arbitrary worker heterogeneity. This approach requires only two inputs: knowledge of steady-state sectoral worker flows over different time horizons and the dispersion of the idiosyncratic shocks. We show that this information is sufficient to construct counterfactual changes in welfare and sectoral

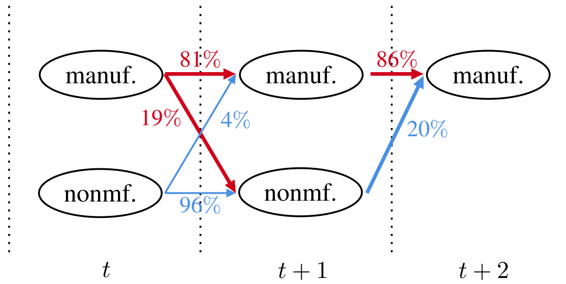


Figure 1. 1-Year and 2-Year Worker Flows in the Data

Notes: The arrows between periods t and $t+1$ represent the steady-state 1-year worker flows between the 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 $\Pr(s_{t+1} = \text{nonmf.}|s_t = \text{nonmf.}) = 1 - \Pr(s_{t+1} = \text{manuf.}|s_t = \text{nonmf.}) = 0.96$. The arrows between periods $t+1$ and $t+2$ provide additional information needed to compute the two-year staying probability for the manufacturing sector. It represents $\Pr(s_{t+2} = \text{manuf.}|s_{t+1} = s_t = \text{manuf.}) = 0.86$ and $\Pr(s_{t+2} = \text{manuf.}|s_{t+1} = \text{nonmf.}, s_t = \text{manuf.}) = 0.20$.

employment in response to sectoral shocks, up to a first-order approximation around a steady state. Steady-state worker flows contain information about the ease with which workers switch sectors over time, and this is precisely what we need to predict how their welfare and sector choices respond to 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 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 how persistent worker heterogeneity shapes our sufficient statistics and the consequences of sectoral shocks. We begin by characterizing the systematic bias in worker flow patterns 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, workers who have self-selected into a sector are more likely to continue choosing 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, we prove that the canonical model always underestimates the welfare losses of adversely affected workers and overestimates the speed of labor reallocation for given exogenous changes in wages. Interestingly, we show that when wages are endogenously determined, the underestimation of welfare losses is likely to be compounded by the overestimation of labor reallocation from given shocks to the labor market.

These findings highlight 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 longitudinal information in the National Longitudinal Survey of Youth 1979 (NLSY79) dataset to compute worker flows over different time horizons.[1] We find that the sectoral worker flows observed in the data are inconsistent with the canonical model without persistent worker heterogeneity. In line with our theoretical finding, the canonical model underestimates the probabilities of workers choosing the same sector by more than a factor of two. The same result holds if we break down workers by their demographic and socioeconomic characteristics, such as gender, education, race, and age. This suggests that the persistent heterogeneity driving this underestimation are mostly 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 one year. The canonical model, which assumes away persistent heterogeneity, necessarily implies that workers who choose to stay in the manufacturing sector between years t and $t+1$ are as likely to stay in the following year as the typical manufacturing worker. However, the data reveals that workers who have self-selected to stay in manufacturing exhibit a higher probability of staying again in the following year ($86\% > 81\%$).[2] 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.[3]

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

[1] Given data constraints, we only compute worker flows over horizons of less than or equal to 18 years.To extrapolate longer-run worker flows from the available data, we leverage the structural model, which is calibrated by matching the observed worker flows.

[2] The data show considerable variation in the frequency of sector switching among workers. Some workers exhibit a high degree of sector mobility, transitioning frequently between different sectors. On the other hand, there are also workers who remain employed in a particular sector for most of their careers and rarely change sectors. This fact alone is difficult to reconcile with the canonical model.

[3] They contribute almost equally to this underestimation. The canonical model implies that the two-year staying probability is $81\% \times 81\% + 19\% \times 4\% = 66\%$, while the actual probability is $81\% \times 5\% + 19\% \times 16\% = 4\% + 3\%$ higher than this value.

sufficient statistics. We put this idea into practice by extending the standard Euler equation approach in the literature to allow for arbitrary worker heterogeneity.

The final part of our paper combines the previous empirical estimates with our 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 manufacturing prices experience a sudden and permanent drop. 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 trade, input-output linkages, and multiple production inputs. We extend the labor supply side of the model by allowing for arbitrary persistent worker heterogeneity.[4] Results demonstrate that when we correctly match the longer-run worker flow patterns with worker heterogeneity, labor reallocation following the China shock is significantly slower, and the negative welfare effects on manufacturing workers are more severe than previously suggested. Without persistent worker heterogeneity, workers initially employed in the manufacturing sector in all states appear to benefit, on average, from the China shock. However, when we do account for persistent worker heterogeneity, the welfare gains of manufacturing workers are close to zero in most states, and manufacturing workers experience welfare losses in five states.

Related Literature. This paper is related to several strands of 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 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 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, building 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 apply this framework to analyze the impact of a trade shock on labor

[4] In contrast to their paper, we make the simplifying assumption that workers can switch sectors only within each state, thus abstracting from interstate migration.

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). Our paper extends the literature by introducing arbitrary persistent worker heterogeneity into the framework and demonstrates how this heterogeneity can significantly influence the results of counterfactual exercises.

In allowing for persistent worker heterogeneity, we relate to a large, mostly static, literature emphasizing 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), 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, allowing us to take into account transitional dynamics and to highlight its importance 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, Arellano and Bonhomme, 2017, Bonhomme, Lamadon, and Manresa, 2022).

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 (with the caveat that not all models in the literature fall within the class of models we consider in this paper). First, our results suggest that structural models can target worker flows over different time horizons in the estimation process 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.

The remainder of the paper is organized as follows. Section 2 presents a model of labor market dynamics with arbitrary persistent 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 bias counterfactual predictions about 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. The final section concludes.

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 persistent (time-invariant) worker heterogeneity. We focus on the dynamic choice of workers over sectors, though the same framework can be applied to analyze their geographic location or occupation choices by a simple relabeling. We first describe the individual worker's dynamic discrete choice problem. The solution to this problem 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 persistent 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 in productivity and nonpecuniary preferences across sectors. At each point in time, workers decide in which sector to work.[5] A worker 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 written as[6]

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

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 log 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 ε_{jt} is worker-specific and reflects nonpecuniary motives for workers to switch sectors.[7] The expectation operator, $\mathbb{E}_t$, is taken over realizations of future wages 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 $C_{ij}^{\omega} = 0$ 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, ε_{jt}, are drawn from a type I extreme-value distribution independently across workers, sectors, and time periods. Let v_{it}^{ω} be the expected value derived from choosing sector i in period t for workers of type ω, taking the average over

[5] Following the timeline of ACM, workers make their sector choice decision one period in advance. This means that the sector choice decision for period t is made in period $t-1$. Once period t arrives, the instantaneous utility w_{it} from their chosen sector is realized, and workers enjoy the realized utility. They then observe the realized value of period t idiosyncratic shocks, $\{\varepsilon_{jt}\}_j$, and subsequently make a sector choice decision for period $t+1$. The expectation operator $\mathbb{E}_t$ is defined with respect to workers' information set at the time of their sector choice decision for period $t+1$.

[6] Given the time series of instantaneous utility, $\{w_{it}^{\omega}\}_{i,t,\omega}$, there are no interactions between workers of different types, so each type of worker solves problem (1) independently.

[7] 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.

the realizations of idiosyncratic shocks $\{\varepsilon_{jt}\}_j$, and let D^{ω}_{ijt} denote 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 sector choice probabilities:

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

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

where the expectation operator $\mathbb{E}_{\varepsilon}$ is taken over the realizations of $\{\varepsilon_{jt}\}_j$. 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. Because there is a continuum of workers for each type ω, we can characterize the law of motion of their sectoral employment share from their sector choice probabilities,

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

We also define the backward transition probability,

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

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 D^{ω}_t and T^{ω}_t, whose (m,n)-element is D^{ω}_{mnt} and T^{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, $\{w^{\omega}_{it}\}$, we can solve for the series of sectoral employment, $\{\ell^{\omega}_{it}\}$, and sectoral values, $\{v^{\omega}_{it}\}$, 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 $\{\mathrm{d}w^{\omega}_{it}\}$ and take first-order approximations of equations (2) and (4) around a steady state.[8]

A *steady state* is associated with time-invariant sectoral wages ($w^{\omega}_{it} = w^{\omega}_{i}$ for all t), where the type-specific value, sector choice 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 D^{ω} and T^{ω}, respectively. The following equations summarize the responses of the endogenous variables in terms of deviations from steady state:

$$\mathrm{d}v^{\omega}_{t} = \mathrm{d}w^{\omega}_{t} + \beta D^{\omega}\,\mathbb{E}_t\,\mathrm{d}v^{\omega}_{t+1}, \tag{5}$$

$$\mathrm{d}\ln\ell^{\omega}_{t+1} = T^{\omega}\,\mathrm{d}\ln\ell^{\omega}_{t} + \tfrac{\beta}{\rho^{\omega}}(I - T^{\omega}D^{\omega})\mathbb{E}_t\,\mathrm{d}v^{\omega}_{t+1}, \tag{6}$$

where we use the vector notation,

$$\mathrm{d}v^{\omega}_{t} = \begin{pmatrix}\mathrm{d}v^{\omega}_{1t} & \cdots & \mathrm{d}v^{\omega}_{St}\end{pmatrix}^{\mathsf{T}},\ \mathrm{d}w^{\omega}_{t} = \begin{pmatrix}\mathrm{d}w^{\omega}_{1t} & \cdots & \mathrm{d}w^{\omega}_{St}\end{pmatrix}^{\mathsf{T}},\ \text{and}\ \mathrm{d}\ln\ell^{\omega}_{t} = \begin{pmatrix}\mathrm{d}\ln\ell^{\omega}_{1t} & \cdots & \mathrm{d}\ln\ell^{\omega}_{St}\end{pmatrix}^{\mathsf{T}}.$$

The algebra follows that of Kleinman, Liu, and Redding (2023), as described in Appendix E. It is worth highlighting the assumptions underlying these equations. Equation (5) is a direct application of the envelope theorem, often referred to the Williams-Daly-Zachary theorem in the discrete choice literature. Thus, it remains valid regardless of the distribution assumption on the idiosyncratic shock.[9] On the other hand, equation (6) is valid only under the assumption that idiosyncratic shock follows a type I extreme-value distribution. Whereas Taylor's theorem implies that it is always possible to express $\mathrm{d}\ln\ell^{\omega}_{t+1}$ as a linear function of $\mathrm{d}\ln\ell^{\omega}_{t}$ and $\mathbb{E}_t\,\mathrm{d}v^{\omega}_{t+1}$ up to the first order, the sufficient statistics results we derive below rely on the coefficients of this linear function's being polynomials of the transition matrices T^{ω} and D^{ω}, which extreme-value distribution assumption guarantees.[10]

It is important to note that equations (5) and (6) 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 a few observable statistics, which can be used to construct counterfactuals. This idea allows us to bypass the well-known challenges associated

[8] We come back to the quality of this first order approximation in Section 5. For the shocks that we consider, the approximation is good.

[9] In fact, the same envelope-type result extends to a much wider class of models, including models with duration dependence mechanism.

[10] This assumption is sufficient but not necessary. In Appendix A.4, we demonstrate that this result is not specific to the extreme-value distribution. Specifically, we show that a version of equation (6) holds for any distribution of the idiosyncratic shock in a limit case of a perturbed economy, in which dispersion of the idiosyncratic shock converges to zero. Thus, all results in this paper apply to this limit case. We also offer alternative primitive assumptions under which coefficients are polynomials of transition matrices.

with explicitly specifying and estimating the distribution of underlying unobserved heterogeneity (e.g., Heckman and Honore, 1990).

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 $(\cdots, \mathrm{d}w^{\omega}_{t-1}, \mathrm{d}w^{\omega}_{t}, \mathrm{d}w^{\omega}_{t+1}, \cdots)$. Lemma 1 summarizes the result.

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

$$\mathrm{d}v^{\omega}_{t} = \sum_{k\geq 0}(\beta D^{\omega})^{k}\mathbb{E}_{t}\,\mathrm{d}w^{\omega}_{t+k}, \tag{7}$$

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

Workers are forward-looking, so all future shocks affect the value of workers and sectoral employment, as can be seen from equations (7) and (8). Due to the presence of switching costs and idiosyncratic shocks, labor reallocation is sluggish, so past shocks also affect sectoral employment, as can be seen from equation (8).

2.3 Welfare and Labor Dynamics at the Macro Level

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

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

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.[11] In so doing, we implicitly assume a utilitarian social welfare function with equal weights across all workers. Next, we define worker flow matrices over different time horizons. As we will demonstrate, these matrices are sufficient statistics for characterizing the welfare and labor market consequences of sectoral shocks.

[11] We therefore abstract from distributional consequences of sectoral shocks across unobservable types ω. Our interest here is in comparing the welfare of workers initially employed in different sectors.

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

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

Unlike the type-specific transition matrices T^ω and D^ω in Lemma 1, these worker flow matrices can be computed directly from longitudinal information on workers' sector choices, as we will do using the NLSY data in Section 4.[12]

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 satisfy either one of the following conditions:*

$$C_{ij}^\omega = C_{ji}^\omega \text{ or } C_{ij}^\omega = \begin{cases} C_i^\omega + \tilde{C}_j^\omega & \textit{if } i \neq j \\ 0 & \textit{if } i = j \end{cases},$$

for all $i, j \in \mathcal{S}$, and $\omega \in \Omega$.

It is worth emphasizing that the first part of Assumption 1 does not require that all workers have the same *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.[13] Similarly, the second part of Assumption 1 does not necessarily imply that all workers have the same labor supply elasticity. 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, T^ω and D^ω, as can be seen in equation (8).[14] Although restrictive, this assumption is standard in the dynamic

[12] In Appendix A.1, we formally show that if we have access to infinite-length longitudinal information on workers' sector choices, we can directly observe the full series of worker flow matrices. However, since panel data have finite time dimension in practice, we can only calculate worker flow matrices $\mathcal{D}_k$ for low enough k's. In Section 4, we discuss various ways of extrapolating worker flow matrices.

[13] In many applications, however, the shocks of interest are known to have heterogeneous effects across *observed* types—for example, between high-skilled and low-skilled workers or across geographic regions. In such cases, we can account for this possibility by conducting the same analysis separately for each observed type of workers. 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.

[14] What this implies is that all heterogeneity in the elasticity of sectoral labor supply with respect to sectoral shocks is revealed by shares. This observation will be useful in deriving our sufficient statistics result.

discrete choice literature, even in papers that incorporate rich heterogeneity of workers. Our approach can be applied to the models studied in those papers.

The conditions in Assumption 2 are often imposed in the literature for various purposes. For example, ACM (in Section IV.D) and Dix-Carneiro (2014) assume that the switching costs can be decomposed as in the second condition to reduce the number of parameters to be estimated. In a different context, Allen and Arkolakis (2014) and Desmet, Nagy, and Rossi-Hansberg (2018) assume the first condition for bilateral trade costs and bilateral switching costs, respectively, in order 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:*

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

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

The logic behind Proposition 1 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, $T^\omega = D^\omega$.[15] Although this equality is not strictly necessary for our purpose, we maintain this assumption in the rest of our analysis. This allows us to derive analytical results in Section 3 and reduces the data requirements needed to implement our sufficient statistics approach.[16] Appendix A.2 shows that these two transition matrices are indeed very similar at the level of granularity at which both matrices can be computed from the data. 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 $(D^\omega)^k$ are aggregated into the k-period worker flow matrix $\mathcal{D}_k$. 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 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.

[15] Two matrices are equal if and only if the steady-state flow of type ω workers from sector i to j is equal to the flow from sector j to i for all sector pairs. However, the definition of the steady state does not necessarily imply this condition, since it only requires that the *total* outflow of type ω workers from a sector be equal to the *total* inflow into that sector. We further need Assumption 2 to guarantee that bilateral worker flows are balanced for all sector pairs.

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

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 $\mathrm{d}v_{t+1}^{\omega}$ in (5) and $\mathrm{d}w_{t+k}^{\omega}$ 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. Relatedly, given the structure in Proposition 1, the aggregate variables do not possess a recursive representation, hence going from equations (5) and (6) to (7) and (8) is key to our results. We will return to this issue in Section 4.

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, $\{\mathcal{D}_k\}$, and the shape parameter ρ. 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.[17] In Section 4, we estimate worker flow matrices and the value of the parameter ρ using panel data. The following corollary summarizes the discussion.[18]

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

We have so far focused on the labor supply side and considered exogenously given wage changes. In Section 2.4, we show that even when we endogenize the wage by augmenting the model with the labor demand side, 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.

[17] 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.

[18] 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.

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 approach effectively accommodates arbitrary worker heterogeneity without encountering the well-known challenges associated with estimating the distribution of unobserved heterogeneity from the data.[19] However, these advantages do not come without costs. First, we need to make restrictions on the set of shocks that can be studied and on the heterogeneity in the dispersion of idiosyncratic shocks (Assumption 1). Second, we rely on the first-order approximation around a steady state, the validity of which depends on the sectoral shock of interest. We will revisit this issue in Section 5.

2.4 Closing the Model: Labor Demand and Equilibrium Wages

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. For this purpose, we need to specify the labor demand side of the model. 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 $\{\varepsilon_{jt}\}_j$ that affect the marginal productivity of labor. The variable ε_{jt} encompasses sector-specific factors, such as capital stock, technology shocks, policy variables, and the like. This relationship can be expressed as $w_{it}^{\omega} = f_i^{\omega}(\{\ell_{jt}\}_j, \{\varepsilon_{jt}\}_j)$.[20] In Appendix A.3, we show that under Assumption 1 we can write this relationship in terms of a first-order approximation as

$$\mathrm{d}w_t = \boldsymbol{A} \cdot \mathrm{d}\ln \ell_t + \boldsymbol{B} \cdot \mathrm{d}\varepsilon_t, \tag{12}$$

[19] Both Dix-Carneiro (2014) and Traiberman (2019) incorporate unobserved heterogeneity in their analyses. However, due to the identification challenge, they are constrained to use a limited number of unobserved types. While this could allow them to capture the absolute advantage of workers, it is difficult to capture the comparative advantage of workers and hence their self-selection into sectors.

[20] For simplicity, we assume that labor demand, unlike labor supply, is determined in a static manner. However, the results below could be extended to models with dynamic labor demand decisions, without affecting any of the main insights.

where there is no ω-index on matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

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_t^ω; type-specific labor allocation across sectors, ℓ_{t+1}^ω; type-specific sector choice probabilities, D_t^ω; 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 $\{v_t^\omega, \ell_{t+1}^\omega, D_t^\omega\}$ 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* $\{\mathrm{d}\varepsilon_t\}$*, 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:*

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

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 prices and wages from the labor demand side to characterize the path of (real) wage changes. The equilibrium is determined as a fixed point of these relations.

Unlike Proposition 2, which requires knowledge of the path of wages, Proposition 1 requires the path of labor demand shocks $\{\varepsilon_i\}$. One can think of it as extending Proposition 2 to the case labor demand that are not perfectly elastic.

Given the shock path, $\{\mathrm{d}\varepsilon_t\}$, 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 value of

ρ, and the reduced-form specification of the labor demand side, $\boldsymbol{A}$ and $\boldsymbol{B}$, the responses of welfare, employment, and wages to labor demand shocks are identical.

3. Employment and Welfare Implications of Persistent Heterogeneity

The canonical dynamic discrete choice model commonly used in the trade, labor, and IO literature abstracts from persistent worker heterogeneity. The sufficient statistics results in the previous section provide a way to account for arbitrary persistent heterogeneity. In this section, we use these results to demonstrate why incorporating this consideration matters in evaluating the consequences of sectoral shocks on welfare and labor reallocation.

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 $\{\mathcal{D}_k\}$.[21] 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 absence of persistent worker heterogeneity imposes on our sufficient statistics; i.e., the worker flow matrices $\{D_k\}$.

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

The first part of Lemma 2 shows that without persistent heterogeneity, the Markovian structure of the model implies that the same transition probabilities are applied to all workers, which enables us to compute the k-period worker flow matrix by multiplying the one-period matrix k times.[22] The second part of Lemma 2 illustrates how accommodating worker heterogeneity relaxes this restriction. With

[21] Another implication of the sufficient statistics result is that, conditional on the observed worker flow matrix series, the results of counterfactual exercises remain unchanged regardless of how we specify the underlying heterogeneity. However, in practice, not all worker flow matrices of the model match those observed in the data, and, as will be seen in this section, models with different worker heterogeneity yield different worker flow matrices.

[22] This is known as the *Chapman-Kolmogorov equation* in the theory of Markovian processes.

worker heterogeneity, 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

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

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

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 (see, for example, Heckman, 1981).

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.[23] 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, 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, Appendix B characterizes the effect of one-time shocks, from which we can calculate the effect of any sequence of shocks.

[23] Even when researchers use panel data that contain the necessary information, it is unclear whether the model correctly matches longer-run worker flows unless they are directly targeted.

Counterfactual Welfare Changes. We begin with welfare changes. 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 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 $\mathcal{D}_k - \mathcal{D}_{k+2}$. For ease of notation, we define b_k as the diagonal element of $\mathcal{D}_k - \mathcal{D}_{k+2}$ that corresponds to sector s:

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

Roughly speaking, b_k measures the rate at which the probability of remaining in sector s decreases 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 $(\mathcal{D}_k)_{s,s}$ 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 faster decline. Assumption 3 further requires the existence of a cutoff $\bar{k}$ at which the order of the speed of decline is reversed. In Appendix B, we show that this assumption does indeed hold with $\bar{k} = 9$ (years) for the worker flow matrix series we observe in the data. Under this assumption, the following proposition shows that the canonical model initially overestimates the change in employment in sector s while underestimating the long-run labor reallocation.

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.[24] In Appendix B, 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. This scenario corresponds, for example, to a small open economy with linear technology affected by changes in world prices induced by trade liberalization. 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. The intuition behind this is straightforward. Without loss of generality, consider a negative shock to a sector. Proposition 4 implies that in response to given negative wage changes, 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 underestimation of the welfare effect.

[24] 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.

In sum, we show that the canonical model, without persistent worker heterogeneity, always 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 inputs to construct counterfactuals for a given shock of interest: the worker flow matrix series, the values of the parameters ρ and β, and the 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 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 the details of the labor demand side.

4.1 Observed Worker Flow Matrices

We compute the worker flow matrices from the National Longitudinal Survey of Youth 1979, a rich dataset compiled by the US Bureau of Labor Statistics.[25] 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.[26] We mainly follow the data-cleaning procedure of Lise and Postel-Vinay (2020).[27] We aggregate sectors into four broad sectors and consider worker flows across these sectors: (i) Agriculture and Construction; (ii) Manufacturing; (iii) Communication and Trade; and (iv) Services and Others.

[25] We also obtain quantitatively and qualitatively similar findings in the monthly Current Population Survey dataset; see Section 5.

[26] This ensures that we consistently measure mobility at a 1-year window.

[27] The survey comprises a cross-sectional subsample that is 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.

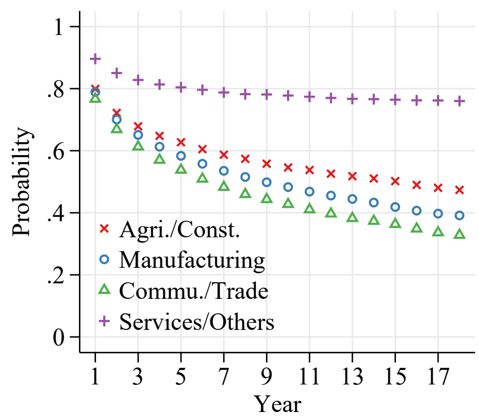


Figure 2. Worker Flow Matrix Series

Notes: Each marker in the figure represents the probability that workers choose the same sector after $k \in \{1, 2, \dots, 18\}$ 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.

Assuming that the economy was in a steady state between years 1980 and 2000, we calculate the series of worker flow matrices by pooling all observations in this period.[28] Given data constraints, we only calculate the k-year worker flow matrices, $\mathcal{D}_k$, up to $k = 18$. Figure 2 plots the diagonal elements of the obtained 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%. The k-period staying probabilities decrease in k, reflecting 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

[28] In Appendix C.1, we show that the results of this section are not driven by the nonstationary nature of the data. Results remain qualitatively and quantitatively similar even when nonstationarity is taken into account.

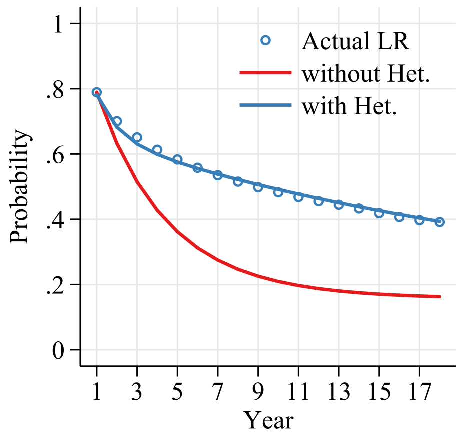


Figure 3. Actual and Model-Implied Worker Flow Matrices: Manufacturing Staying Probabilities

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

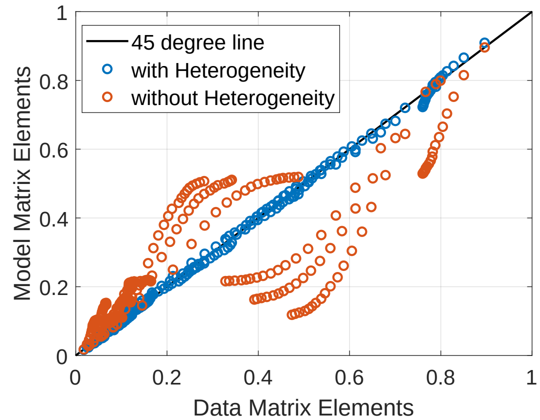


Figure 4. Actual and Model-Implied Worker Flow Matrices: All 4×4 Elements

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

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.[29] Figure A.4 further shows that this underestimation is not driven by the nonstationary nature of the data.

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 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 corresponding 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, along with the

[29] 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 worker flow matrices for two periods $t = 1, 2$ are D^1 and D^2, respectively. Under the stationarity assumption, we calculate the worker flow matrix $\bar{D}$, which is applied to both periods, by taking an average of D^1 and D^2. Suppose that we put equal weights on D_1 and D_2. Then, we have

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

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.

45-degree line. The points clustered below the 45 degree line correspond to the diagonal elements of the worker flow matrices that 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 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 three dimensions, gender, race, and education, which 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, and low-skilled (less than high school and high school) and high-skilled (some college and college or more). Unique combinations of these three dimensions of heterogeneity define eight worker types. In Figure A.5, we plot the actual manufacturing staying probabilities separately for these eight worker types. Indeed, workers who differ along these characteristics exhibit highly distinct sectoral movement patterns. 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 three dimensions of observed characteristics. Specifically, we consider a model with the eight observed worker types. For each worker type ω^{obs}, we can calculate the 1-year transition matrix $(D^{\omega^{\text{obs}}})$ directly from the data. Since workers are assumed to be homogeneous within each of the eight types, their k-year transition matrix can then be computed by $(D^{\omega^{\text{obs}}})^k$. 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 solid 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

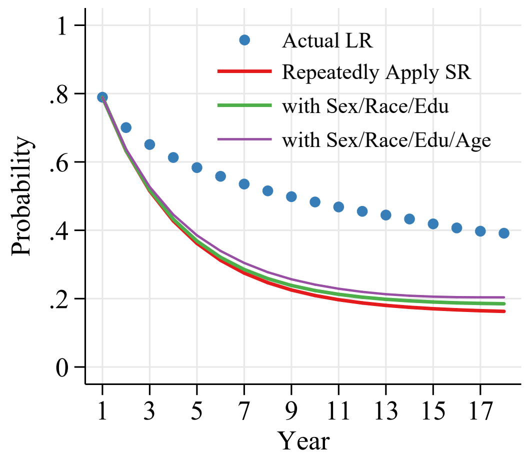


Figure 5. Predicted and Actual Staying Probabilities, with and without Socioeconomic Characteristics

Notes: The blue dots in the figure represent k-period manufacturing staying probabilities. The blue solid line represents the fit of the estimated two-type worker model. The red solid line represents the fit of the estimated canonical model. The green line incorporates three observed characteristics. The purple line additionally incorporates age. Data source: NLSY79.

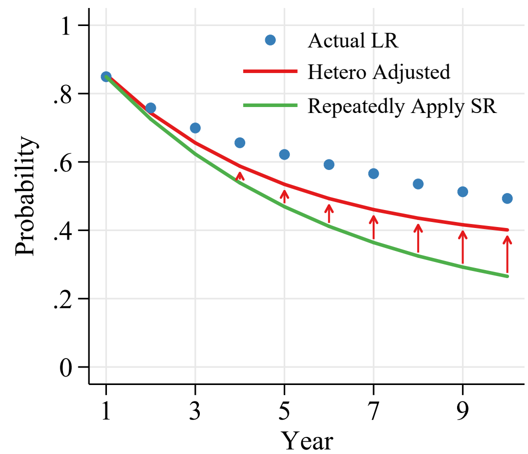


Figure 6. Predicted and Actual Staying Probabilities, with and without Socioeconomic Characteristics

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

actual pattern of worker flows. If we also incorporate age in the model, there is slight improvement in the fit, but the implied worker flow matrix series still significantly differs from the actual one (purple solid line).[30] In sum, workers with different demographic and socioeconomic characteristics do indeed behave differently, but this fact barely changes *aggregate* labor market dynamics.[31] In turn, the sufficient statistics result implies that adding these worker characteristics into a model does not cause substantial changes to the results of the counterfactual analysis.[32]

Pure Duration Dependence. The limited role of demographic and socioeconomic characteristics implies two possibilities: either these characteristics are just poor proxies for underlying worker heterogeneity, or the gap in Figure 3 arises from mechanisms other than worker heterogeneity, such as

[30] 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, "old" and "young," so that each represents about 50% of the total sample.

[31] Similar results can be found in various fields of economics. For example, Card, Rothstein, and Yi (2023) find that only a modest fraction of the variation in average wages across commuting zones is explained by differences in the observed characteristics of workers.

[32] This means that the discrepancy in Figure 3 is mostly driven by within-type heterogeneity instead of across-type heterogeneity. In Appendix C.2, we demonstrate that we can observe similar gaps between actual and predicted longer-run staying probabilities for each of the eight types. In particular, the gaps are more pronounced for female Hispanic/Black workers.

duration dependence. A notable example is the accumulation of sector-specific human capital, where otherwise identical workers who have spent more time in manufacturing may have accumulated more manufacturing-specific human capital, rendering them more likely to choose the sector again. Other examples include fixed adjustment costs (e.g., Stokey, 2008), learning about match productivity (e.g., Jovanovic, 1979), and psychological choice models (e.g., Cain, 1976).[33] This raises a concern, given that only the first equation of Proposition 1 extends to the case with a duration dependence mechanism.

In response to this issue, we provide suggestive evidence pointing to the importance of worker heterogeneity in generating the gap 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. Differences among workers materialize as differences in their sector choice patterns, so their histories allow us to effectively control for their heterogeneity. Specifically, we use pre-1990 data to compute the 1-year manufacturing staying probability for each worker, then categorize workers into five types based on the quintiles of this probability. We then 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 shown as blue dots, while the green and red lines represent the staying probabilities implied by the canonical model and the model with five worker types, respectively. Note that the red line is computed under the assumption that workers are homogeneous within each of the five worker types and that there is no duration dependence mechanism in play. The red line is close to blue dots, and the gap between the actual and implied staying probabilities is reduced by more than half. Given that worker heterogeneity is only partially controlled by sector choice history, this result suggests that at least half of the gap is due to worker heterogeneity, the mechanism emphasized in this paper.

4.4 Extrapolation Using the Structural Model

To apply sufficient statistics results, we need 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. Leveraging the structural model provides a natural method for extrapolation.[34] For

[33] 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. Alvarez, Borovičková, and Shimer (2016) also study how to distinguish between the two in the context of unemployment duration.

[34] 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 Appendix C.6, we explore alternative extrapolation methods that do not rely on a structural model. In addition to extrapolation, there are two other benefits of the structural

Table 1: Estimation Results

$\lvert\Omega\rvert = 2$	Type ω_1 (31.1%)					Type ω_2 (68.9%)				
Sector	Wage	Switching Cost				Wage	Switching Cost			
Agri/Const	0.58	0.00	1.53	1.69	1.55	1.02	0.00	4.62	5.62	5.46
Manufacturing	0.60	1.53	0.00	1.33	1.41	1.02	4.62	0.00	4.88	4.93
Commu/Trade	0.66	1.69	1.33	0.00	0.98	1.00	5.62	4.88	0.00	3.72
Services/Others	0.77	1.55	1.41	0.98	0.00	1.07	5.46	4.93	3.72	0.00

this purpose, we estimate the structural model by matching the worker flow matrices we computed directly from the NLSY data, $\{\mathcal{D}_k\}_{k=1}^{18}$. The estimated structural model generates the full set of worker flow matrices, which are used in Section 5 to perform counterfactual exercises.[35] 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.

Specifically, we estimate the number of worker types along with their respective steady-state instantaneous utility vectors w_i^ω and switching costs C_{ij}^ω by matching 18 worker flow matrices (i.e., 216 moments). Note from the worker's sector-choice problem (1) that only the ratio between these values and the parameter ρ can be identified from the observed worker flow matrices. Thus, we only estimate these ratios until we estimate the value of ρ in Section 4.5. Following Assumption 2, we impose symmetry on the switching costs. The estimation process involves two steps: We first maximize the likelihood of observing $\{\mathcal{D}_k\}_{k=1}^{18}$ to estimate $\{\frac{1}{\rho}w_i^\omega, \frac{1}{\rho}C_{ij}^\omega\}$ for a given number of worker types, then use the Bayesian information criterion to determine the number of worker types; see Appendix C.5 for more details.

Table 1 shows the estimation result. The Bayesian information criterion supports the model with two worker types. Figure 7 plots the resulting transition matrix for each type of worker. The first type has a comparative advantage in non-manufacturing sectors and low switching costs. Thus, workers of this type switch sectors frequently, as indicated by the small diagonal elements of the transition matrix in Figure 7. In contrast, the second type has much higher switching costs, so workers of this type rarely move to other sectors. In Appendix C.11, we show how to interpret the figures in Table 1. In particular, paying one unit of switching costs means paying 3.25% of lifetime consumption. Thus, the switching costs in Table 1 are at most less than 20% of lifetime consumption. This is smaller than the estimates of

estimation. First, it serves as a proof of concept: By seeing whether our structural model can match the observed worker flow matrix series, we can test whether our framework is consistent with the data. Second, we can use the estimated structural model to evaluate the quality of the first-order approximation around a steady state. In Section 5, we compare the results of counterfactual exercises computed from the exact solution of the estimated model and those computed using the sufficient statistics result.

[35] In fact, we can simply treat the estimated model as if it were the true model and use it to perform counterfactual exercises directly. The sufficient statistics result guarantees that we will obtain correct counterfactual results regardless of whether this model is true.

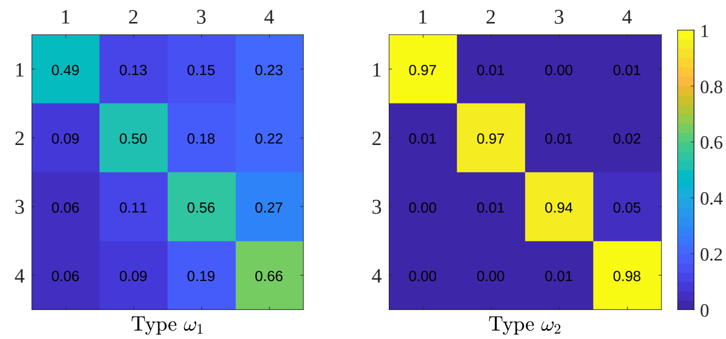


Figure 7. Type-Specific Transition Matrix

ACM, who find that the *average* switching cost is at least 20% of lifetime consumption. Our estimates are close to those of Artuç and McLaren (2015), in which the switching costs are distributed around 12% of lifetime consumption.[36]

In Appendix C.3, we also estimate the canonical model by matching the one-period worker flow matrix, $\mathcal{D}_1$. All parameters, including elements of the transition matrix, lie between the corresponding parameters for the model with two worker types.

Model Fit. 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. The figure shows that the model with only two worker types closely matches actual short-run and longer-run staying probabilities. We obtain similar results for the other elements of the worker flow matrices. In Figure 4 we use blue dots to plot all 4×4 elements of the model-implied worker flow matrix series against the actual value in the data, along with the 45-degree line. Most of the dots lie roughly on the 45-degree line. The fit of the model is surprising for two reasons. First, the degrees of freedom (19 parameters) are much smaller than the number of moments we target (216 moments).[37] Second, the dynamic discrete choice framework imposes systematic restrictions on the model-implied worker flow matrices, so we could not match any worker flow matrix series even with an arbitrary number of worker types.

[36] As Artuç and McLaren (2015) argue, one reason for the smaller estimates is the inclusion of sector-specific nonpecuniary benefits in the model, which are absent in ACM's model.

[37] Suppose we want to match only the 1-year worker flow matrix. We can perfectly match this matrix with only one type of worker if we can choose an arbitrary transition matrix for this type. In terms of degrees of freedom, we match $N(N-1)$ values with $N(N-1)$ parameters, where $N = 4$ is the number of sectors. Suppose we also want to match one more worker flow matrix. This exercise can be seen as matching the level and slope of the dots in Figure 2. At least in terms of degrees of freedom, we can achieve this with only two types of workers: matching $2N(N-1)$ values with $2N(N-1)+1$ parameters ($N(N-1)$ parameters for each transition matrix, and 1 for the type share). However, we also want to match the overall shape of the dots in Figure 2, and confine ourselves to the case in which the type-specific transition matrices are generated by the structural primitives. Thus, we end up with 19 parameters that can be used to match 216 moments.

The flexibility due to worker heterogeneity, characterized in Lemma 2, is necessary to match the observed worker flow matrices. As seen in Section 4.2, the canonical model with one type of workers fails to match the observed worker flow matrices. Table 1 clearly reveals why the canonical model significantly underestimates longer-run staying probabilities. Workers of the second type rarely change sectors and have comparative advantage in manufacturing. Thus, conditioning on the fact that workers have previously self-selected into the manufacturing sector greatly increases the probability that they are the second type, and thus increases the probability that they will stay or choose again the manufacturing sector in subsequent periods. While the fit of the model improves substantially with two types of workers, the additional increase in fit from adding more types of workers is negligible, causing the Bayesian information criterion to choose the two-type worker model. See Figure A.7 for comparison of the fits of the two-type model and the five-type model.

Identification. Comparing the fits of the models does not necessarily identify the true number of worker types, let alone the fact that two worker types cannot capture the multifaceted nature of real-world worker differences. However, a key feature of our approach is that it does not require identification of all elements of the true model. As long as the estimated model closely approximates the observed worker flow matrix series, our sufficient statistics result ensures that the model always provides valid counterfactuals for the outcome of interest—namely, aggregate welfare and employment. This feature distinguishes our approach from the latent variable approach in the literature, such as the finite mixture model and k-means clustering (e.g., Arcidiacono and Jones, 2003; Heckman and Singer, 1984; Bonhomme, Lamadon, and Manresa, 2022), in which the validity of counterfactual predictions requires a higher level of confidence in identification.

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.[38] The possibility of such an estimation is already suggested in the second

[38] In Propositions 1 and 2, we showed that the worker flow matrix series serves as sufficient statistics for counterfactual exercises *when we know the value of the parameter ρ*. However, if the estimation procedure for ρ depends on how we specify worker heterogeneity, the distribution of worker heterogeneity can affect the results of counterfactual exercises through its effect on estimation of ρ. In this sense, our estimation method is comparable to estimation of the trade elasticity using the gravity equation of Arkolakis, Costinot, and Rodríguez-Clare (2012). Just as their estimation based on the gravity equation provides a valid estimate within a class of "quantitative trade models," our method yields a valid estimate within a specific class of dynamic discrete choice models.

equation of Proposition 1, which we restate here for convenience:

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

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.[39]

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.[40]

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 $\mathcal{D}_k = (\mathcal{D}_1)^k$ to rewrite equation (15) recursively as follows:[41]

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

See Appendix C.7 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, in Appendix C.7, we show that when there is a finite number of worker types, N, equation (15)

[39] The responsiveness is inversely related to the value of ρ. A higher value of ρ indicates that sector choice decisions are primarily driven by idiosyncratic shocks. As a consequence, the impact of wage shocks on sectoral employment is relatively diminished when ρ is high.

[40] ACM (p.1040) write that "Perhaps the biggest weakness in the Euler-equation approach ... is that it assumes away workers with unobserved [heterogeneity] .... A full exploration of such effects probably requires a structural micro approach." However, the method described in this section demonstrates that up to the first-order approximation, this problem can be circumvented.

[41] 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 equations that involve migration probabilities rather than labor supply.

always possesses an *exact* recursive representation of the form[42]

$$\mathrm{d}\ln\ell_t = \sum_{k=1}^{4N-2} \mathbf{\Gamma}_k \mathbb{E}_t\, \mathrm{d}\ln\ell_{t+k} + \tfrac{\beta}{\rho} \sum_{k=1}^{4N-3} \mathbf{\Lambda}_k \mathbb{E}_t\, \mathrm{d}w_{t+k} \tag{17}$$

where $\mathbf{\Gamma}_k$ and $\mathbf{\Lambda}_k$ are functions of worker flow matrix series, $\{\mathcal{D}_k\}$. Second, recall from Section 4.4 that a model with two worker types provides a close approximation to the observed worker flow matrix series $\{\mathcal{D}_k\}$ in the data. Combining these two findings, we can conclude that equation (15), with $\{\mathcal{D}_k\}$ from the data, can be approximately represented in the recursive form (17) with $N = 2$. In Appendix C.8, we provide formulas for computing $\{\mathbf{\Gamma}_k\}_{k=1,\ldots,6}$ and $\{\mathbf{\Lambda}_k\}_{k=1,\ldots,5}$. In Appendix C.10, we use simulation to demonstrate that the obtained approximate recursive representation provides a close fit to the actual dynamics of sectoral employment.

We further modify the obtained recursive representation in two ways:

$$\ln\ell_t - \sum_{k=1}^{6} \mathbf{\Gamma}_k \ln\ell_{t+k} = \tfrac{\beta}{\rho} \sum_{k=1}^{5} \mathbf{\Lambda}_k w_{t+k} + \mathsf{ExpectationError}_{t+1,t+6}\,. \tag{18}$$

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 $\sum_{k=1}^{6} \mathbf{\Gamma}_k = I$ and $\sum_{k=1}^{5} \mathbf{\Lambda}_k = \boldsymbol{O}$, where I is the identity matrix and $\boldsymbol{O}$ 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, $\sum_{k=1}^{5} \mathbf{\Lambda}_k\, \mathrm{d}w_{t+k}$, 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.

[42] For further insights into how this result relates to the findings of Granger and Morris (1976), see Appendix C.7.

Table 2: 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 $^{***}: p < 0.01$. Data source is NLSY79 and BLS.

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.[43] 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 *log* sectoral wage.

Table 2 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 Appendix C.11). Comparing the results in Columns (1) and (2), it appears that the estimated coefficient

[43] The assumption of a four-sector model raises an econometric issue similar to the one discussed by ACM. If the true model consists of more than four sectors with a dispersion parameter ρ, is it valid to approximate the model with a four-sector model and estimate the dispersion parameter based on this approximation? More importantly, can we use it to conduct counterfactual exercises? In Appendix C.9, we demonstrate that under certain assumptions on switching costs, the validity of using the approximated model is supported by the fact that the maximum of i.i.d. type I extreme-value distributed random variables follows another type I extreme-value distribution with the same scale parameter.

$\widehat{\beta/\rho}$ from OLS 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.[44] 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 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 $\rho = 0.56$.[45] This estimate in Column (2) is our preferred specification, and we will use it in the subsequent analysis. However, our purpose is to see how the results of the counterfactual exercises would change if we accounted for worker heterogeneity—not to see how different values of ρ would change the results. Thus, we will also report the results of counterfactual exercises under $\rho = 0.5$ and $\rho = 2$ in the appendix.

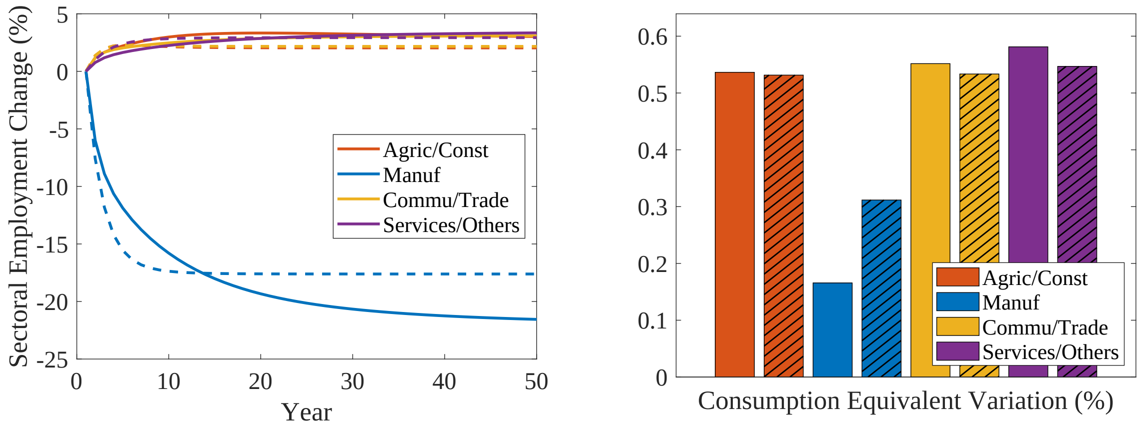
We now move on to the application of our model, in which we use the estimated model with worker heterogeneity to quantify the effect of sectoral shocks.

5. Applications

In this section, we consider two different sectoral shocks and quantify the implication of persistent worker heterogeneity for welfare and labor reallocation. We first apply our results to a stylized trade liberalization exercise of ACM, which illustrates how the framework of the literature can easily be extended to allow for worker heterogeneity. For a more realistic application, we then examine the dynamic effects of the rise in China's import competition on the 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.

[44] 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.

[45] 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 8. Counterfactual Changes in Sectoral Employment and Welfare: Trade Liberalization

Notes: These figures plot the transitional dynamics following an unexpected permanent drop in manufacturing prices. 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.

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 the 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 Appendix D.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 perfect-foresight transition path following announcement of the shock in year 1 until the economy reaches a new steady state.

Figure 8 shows results of the counterfactual exercise. In Figure 8a, 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.[46]

In Figure 8b, we plot changes in welfare, measured in terms of consumption-equivalent variation; that is, the proportional change in lifetime consumption sequence that would have the same effect on household welfare as would the welfare effect of the China shock. In particular, plain bars represent the welfare changes of workers based on their initial sector of employment predicted by the model. Hatched bars represent the 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.[47] However, consistent with the prediction of Proposition 3, the canonical model overestimates the gains of manufacturing workers by nearly a factor of two. At the same time, it slightly underestimates the gains of non-manufacturing sector workers, resulting in a 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 in Appendix D.1 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.

To gain a 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 Appendix D.1.[48] When the shock hits the economy, type 1 workers, who have lower switching costs and comparative advantages in non-manufacturing sectors, can easily move out of manufacturing and

[46] In Appendix D.1, we compute the impulse responses of sectoral employment from the observed worker flow matrix series. As predicted by Proposition A.3, 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 8a.

[47] Appendix D.1.2 plots the time path of changes in sectoral wages. 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, manufacturing employment share declines over time and in the new steady state.

[48] An important caveat to this interpretation pertains to the identification of the structural model. In estimating the structural model, many configurations of the model primitives are almost equally successful in matching the worker flow matrix series. This means that even small changes in the observed matrices can lead to significantly different estimation results. Thus, statements made at the unobserved type level should be viewed with caution, as they may be far from the true world.

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 A.10, 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. Despite the relatively large magnitude of the 10% shock considered in this section, the sufficient statistics results deliver a close approximation. In particular, the approximation error is an order of magnitude smaller than the difference between the counterfactual predictions of the models with and without worker heterogeneity. The 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 welfare of US workers and labor reallocation. We closely follow the dynamic quantitative trade model of Caliendo, Dvorkin, and Parro (2019) (hereafter, CDP) and extend the labor supply side by allowing for arbitrary persistent worker heterogeneity.[49] We make two simplifying assumptions relative to the original specification of CDP. First, we consider 4 broad sectors—Manufacturing, Wholesale/Retail, Construction, and Services—and another auxiliary sector representing nonemployment.[50] Second, for reasons discussed shortly, we abstract from interstate migration and assume that workers can switch

[49] 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); and learning and expectations (Fan, Hong, and Parro, 2023). Incorporating worker heterogeneity in models with these additional features is an important direction for future research.

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

sectors only within each state.[51,52] 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. In this exercise, 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 the 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, as shown in Figure A.12, 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 the state level, we can account for such state-level heterogeneity. With our simplifications, 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 $\mathcal{D}_k^{\text{state}}$ 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.[53,54] To compute worker flow matrices for other values of k, we follow Section 4 and estimate the structural model with two types of workers that best match the observed worker flow matrices. We then use the estimated model to extrapolate longer-run worker flow

[51] 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.

[52] In principle, allowing for regional migration would dampen the welfare effects because it provides an additional margin of adjustment. However, US workers change sectors nearly 10 to 100 times more often than they change states, suggesting 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 to their model's prediction. Autor, Dorn, and Hanson (2013) and Autor et al. (2014) also demonstrate that regional migration is not an important mechanism through which the economy adjusts to the China shock.

[53] In contrast to the literature, which often uses only one worker flow matrix—either monthly ($k = 1$) or quarterly ($k = 3$)—we exploit the information contained in all worker flow matrices to identify underlying persistent worker heterogeneity. For any k, if we use the k-period worker flow matrix $\mathcal{D}_k^{\text{state}}$ and rely on the homogeneous-worker assumption to calculate the remaining worker flow matrices, we would underestimate the diagonal elements of k'-period worker flow matrices for k' greater than k.

[54] 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 $\mathcal{D}_k^{\text{state}}$ 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 A.13, 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 from sample attrition and the resulting selection (e.g., Moscarini and Thomsson, 2007).

matrix series.[55] In Figure A.14, 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.[56]

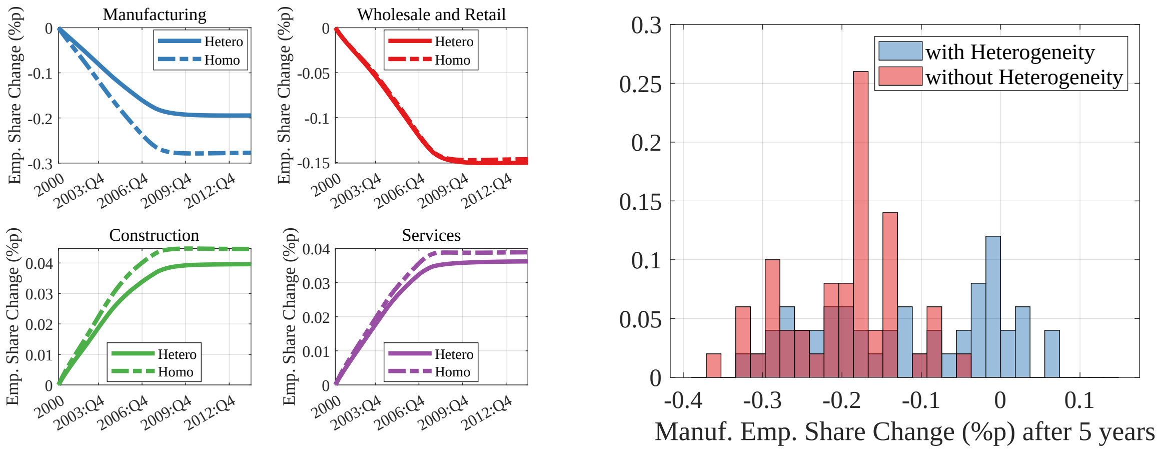
Labor Demand Side and the China Shock. The labor demand side of the model is more complex than in Section 5.1: It features a large number of labor markets distinguished by sector and region, international trade, interregional trade within the US, input-output linkages, and multiple production inputs. Appendix D.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 Appendix D.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 as 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 the 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 Appendix D.1 for the 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 statistics result; see Appendix D.2 for details.[57]

[55] One concern is that this requires too much extrapolation. In Figure A.16, 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.

[56] The estimated ρ is at the opposite extreme of the CDP's estimate within the range of estimates in the literature. Thus, we also report results under the CDP's estimate in Section 4.5. Another issue is that we estimate the parameter ρ at an annual frequency. In Appendix D.2, we present a way to transform this to quarterly frequency. The resulting value is given by $\rho = 1.0011$.

[57] 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 advantage, 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).



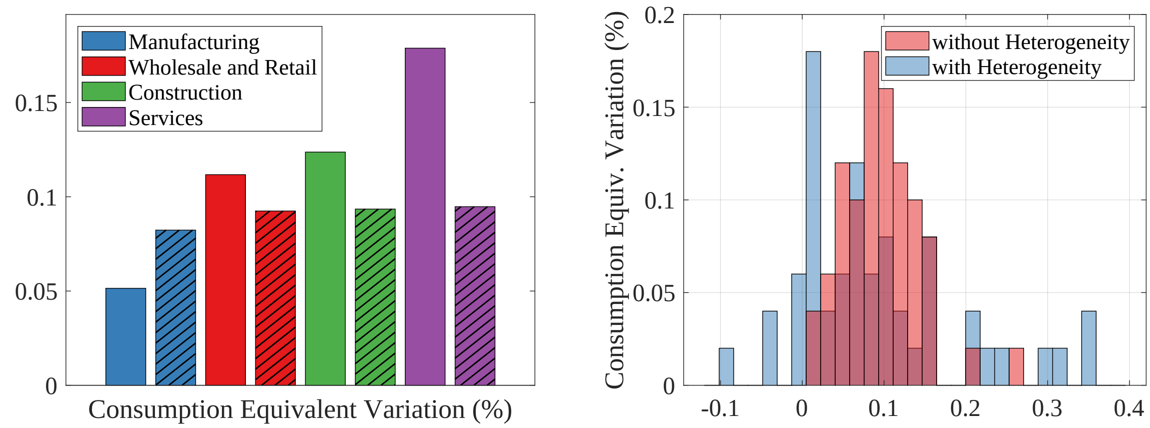
(a) Manufacturing Employment: Aggregate Effect

(b) Manufacturing Employment: Individual States

Figure 9. Effect of the China Shock on Welfare

Figures 9 and 10 plot results of counterfactual exercises computed using the sufficient statistics approach. In Figure A.18, we also compare these results with those calculated from the exact solution of the model. The small differences between them alleviate concerns regarding the first-order approximation around a steady state in the context of the China shock, arising from the downward secular trend in US manufacturing.

Sectoral Employment Changes. Figure 9a 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 manufacturing production from the US to China, resulting in a decline in the share of US manufacturing employment (top left panel). With worker heterogeneity, the China shock reduces the share of manufacturing employment by around 0.2 percentage points (or, equivalently, 0.45 million manufacturing jobs) after 10 years. At the same time, the China shock increases employment in the construction and services sectors, as observed in the data. These sectors benefit from access to cheaper intermediate goods made available by the China shock. Dashed lines plot the 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 50%. Figure 9b presents a histogram of state-level changes in manufacturing employment after 5 years. The impact of the China shock varies across states in both models, but in line with Figure 9a, the contraction of manufacturing is faster in the canonical model.



(a) Sectoral Welfare Changes: Aggregate Effect

(b) Manufacturing Welfare Changes: Individual States

Figure 10. Effect of the China Shock on Sectoral Employment

Welfare Changes. In terms of welfare, the heterogeneous-worker model predicts a 0.09% increase (in terms of consumption-equivalent variation), which is similar to the 0.11% 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 win and the losers lose. In Figure 10a, we plot sectoral welfare changes measured at the time the China shock was known to the US labor market. Again, plain bars represent the sectoral welfare changes predicted by the heterogeneous-worker model, and hatched bars represent rhose 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, 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 half of the gains of workers in the other sectors.

Similar to employment effects, the welfare impact of China's import competition varies substantially across regions. In Figure 10b, 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 5 states—Alabama, Illinois, Louisiana, Michigan, and Ohio—experience a decline in welfare.[58] These states have the highest manufacturing employment shares, experiencing 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.

Figure A.19 presents the percentage changes in sectoral wages averaged across states induced by the China shock. The average effect is positive for all sectors, although some states experience declines in the manufacturing real wage. 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, in Figure A.20 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. The result implies that around a quarter of the welfare change gap between the two models arises from this difference in real wage changes, and the remaining three-quarters result from differences in the worker flow matrix series. This highlights the importance of endogenizing wage changes when studying the role of worker heterogeneity.

In sum, results of the counterfactual exercises demonstrate the quantitative importance of accounting for worker heterogeneity.

[58] Workers from these five states continue to experience negative welfare effects even in the long run, which is in stark contrast to the findings in the literature.

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, 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.

Alvarez, Fernando E, Katarína Borovičková, and Robert Shimer (2016) "Decomposing Duration Dependence in a Stopping Time Model,"Technical report, National Bureau of Economic Research.

Arcidiacono, Peter, and John Bailey Jones (2003) "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica*, 71 (3), 933–946.

Arellano, Manuel, and Stéphane Bonhomme (2017) "Nonlinear Panel Data Methods for Dynamic Heterogeneous Agent Models," *Annual Review of Economics*, 9, 471–496.

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.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2022) "Discretizing Unobserved Heterogeneity," *Econometrica*, 90 (2), 625–643.

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.

Cain, Glen G (1976) "The Challenge of Segmented Labor Market Theories To Orthodox Theory: A Survey," *Journal of economic literature*, 14 (4), 1215–1257.

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.

Card, David, Jesse Rothstein, and Moises Yi (2023) "Location, Location, Location," Working Paper 31587, National Bureau of Economic Research.

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.

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

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.

Granger, Clive WJ, and Michael J Morris (1976) "Time Series Modelling and Interpretation," *Journal of the Royal Statistical Society: Series A (General)*, 139 (2), 246–257.

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.

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.

Jovanovic, Boyan (1979) "Job Matching and the Theory of Turnover," *Journal of political economy*, 87 (5, Part 1), 972–990.

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.

Kim, Ryan, and Jonathan Vogel (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, 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*.

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

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.

Stokey, Nancy L (2008) *The Economics of Inaction: Stochastic Control Models with Fixed Costs*: Princeton University Press.

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

Appendix

Appendix will be added soon. Please click here for the latest version.