Online Appendix for
Sectoral Shocks and Labor Market Dynamics: A Sufficient Statistics Approach

Jaeeun Seo Ryungha Oh

December 2024

OA.1 Theoretical Results

Throughout this section, we assume that the inequality in Lemma 2 holds—i.e., either assume that worker flow matrices are as in the data or that there are two sectors, $S = 2$.

OA.1.1 Single Crossing Condition

We know from Lemma 2 that, compared with the canonical model calibrated by matching the one-period worker flow matrix, the model with worker heterogeneity implies lower values of b_k at least for $k = 0$ and $k = 1$:

$$b_0 = \left(\mathcal{F}_0 - \mathcal{F}_2\right)_{ss} = 1 - (\mathcal{F}_2)_{ss},$$
$$b_1 = \left(\mathcal{F}_1 - \mathcal{F}_3\right)_{ss} = (\mathcal{F}_1)_{ss} - (\mathcal{F}_3)_{ss}.$$

Thus, $\bar{k} \geq 1$. In Figure OA.1, we plot the difference between $\{b_k\}$ implied by the canonical model and those observed in the data (extrapolated using the method described in the main text). Initially, the canonical model yields larger values of b_k, but eventually it leads to smaller values compared with those implied by the data (although not plotted, this is true for all values of k greater than 9; i.e., there is no more crossing).

OA.1.2 Response to a One-time Shock

Consider a one-time negative shock to a sector $s \in \mathcal{S}$ that is known to agents in period 1:

$$\mathrm{d}w_{s\tau} = -\Delta < 0 \tag{OA.1}$$

for given $\tau > 1$. The effect of any series of negative shocks to sector s can be calculated as the sum of the effects of such one-time negative shocks.

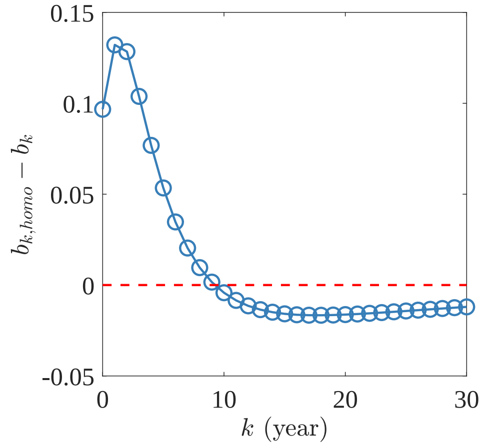


Figure OA.1. Differences in b_k Series: Manufacturing Sector

Notes: For each k, this figure plots the difference between the diagonal elements of $(\mathcal{F}_k - \mathcal{F}_{k+2})$ corresponding to the manufacturing sector implied by the canonical model and those observed in the data. For the canonical model, we compute $\mathcal{F}_k$ by multiplying $\mathcal{F}_1$ k times. Since we only observe a finite number of worker flow matrices in the data, we extrapolate it using the estimated structural model. The same pattern is observed for the other sectors. Data source: NLSY79.

Effects on Sectoral Welfare. With worker heterogeneity, workers initially employed in sector s are more likely to stay in sector s when the shock hits the sector. Thus, they suffer more from the one-time shock.

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

This result in turn implies that for any series of negative shocks to sector s, the canonical model underestimates the negative welfare effect on workers initially employed in sector s, proving Proposition 3.

Effects on Sectoral Employment. The following proposition characterizes the condition under which the canonical model overestimates the decline in employment in sector s in period $t > 1$.

Proposition OA.2. *Consider a one-time negative shock to sector s of the form* (OA.1) *known to agents in period* 1. *Under Assumption 3, there exists a decreasing function* $B : \mathbb{N} \to \mathbb{N}$ *such that the canonical model overestimates the decline in employment in sector s in period t in response to the shock if and only if* $|t - \tau| \leq B(t \wedge \tau)$*, where* $t \wedge \tau$ *denotes the minimum of t and* τ.

When $|t - \tau|$ and/or $t \wedge \tau$ are small, the canonical model calibrated by matching the one-period worker flow matrix overestimates the decline in employment in sector s in period t in response to the shock, $\frac{\partial \ln \ell_{s,t}}{\partial w_{s,\tau}}$.[1]

The result implies that whether the models without worker heterogeneity overestimate or underestimate 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 overestimation of the decline in

[1] If w is log wage, this measures the elasticity of sectoral employment with respect to sectoral wages.

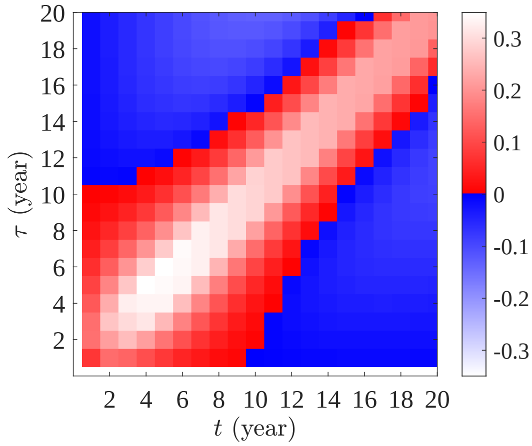


Figure OA.2. Differences in the Response of Sectoral Employment

Notes: This figure plots the values of $\frac{\partial \ln \ell_{s,t}}{\partial w_{s,\tau}}\Big|_{\text{canonical}} - \frac{\partial \ln \ell_{s,t}}{\partial w_{s,\tau}}\Big|_{\text{data}}$ for $1 \leq t, \tau \leq 20$. These derivatives are calculated from equation (11) using the worker flow matrices $\mathcal{F}_k$ from the NLSY79 data for $\frac{\partial \ln \ell_{s,t}}{\partial w_{s,\tau}}\Big|_{\text{data}}$ and using the worker flow matrices implied by the canonical model for $\frac{\partial \ln \ell_{s,t}}{\partial w_{s,\tau}}\Big|_{\text{canonical}}$.

employment in a negatively affected sector. This intuition is what Proposition OA.2 describes when $|t - \tau|$ and/or $t \wedge \tau$ are small; i.e., when the shock is recently known or when the period affected by the shock is close to the period of interest. On the other hand, in the canonical model, workers have relatively lower probabilities of remaining in a sector, which implies that their sector choices in a given period do not have long-lasting impacts on their future sector choices. This aspect works in the opposite direction to our previous intuition and can become dominant if $|t - \tau|$ and/or $t \wedge \tau$ are large enough. For example, suppose that t is much larger than τ. A negative shock to a sector's wage in period τ reduces employment in the sector around that period. However, this reduced employment has only a limited impact on employment in the sector in the distant period t in the canonical model. Figure OA.2 plots the relative size of the decline in employment of sector s in period t in response to a shock to the period τ wage predicted by the canonical model and that implied by the data. As expected, the decline is overestimated in the canonical model for small t or τ or small $|t - \tau|$ (red cells). In short, the canonical model tends to overestimate the short-term impact of shocks on sectoral employment but underestimates their long-term effects. In particular, within a 7-year time horizon, the canonical model consistently overestimates the impact of shocks on sectoral employment.

OA.1.3 Two-Sector Model

In this section we consider a special case of our model with two sectors, $\mathcal{S} = \{1, 2\}$. The results in this section will be used for the proof of Lemma 2.

For simplicity, we assume that there are a finite number of worker types. These types are indexed by ω, and the population share of type ω is $\theta_\omega \in (0,1)$. We denote the transition matrix of type-ω workers by

$$F_\omega = \begin{pmatrix} \bar{\alpha}_\omega & \alpha_\omega \\ \beta_\omega & \bar{\beta}_\omega \end{pmatrix},$$

where $\bar{\alpha}_\omega = 1 - \alpha_\omega$ and $\bar{\beta}_\omega = 1 - \beta_\omega$.

By induction, we can obtain a general formula for the elements of the matrix F_ω^k.

Lemma OA.1. *F_ω^k has the following form:*

$$F_\omega^k = \begin{pmatrix} 1 - \alpha_\omega f^k(\bar{\alpha}_\omega + \bar{\beta}_\omega) & \alpha_\omega f^k(\bar{\alpha}_\omega + \bar{\beta}_\omega) \\ \beta_\omega f^k(\bar{\alpha}_\omega + \bar{\beta}_\omega) & 1 - \beta_\omega f^k(\bar{\alpha}_\omega + \bar{\beta}_\omega) \end{pmatrix},$$

where $f^k(x) = \frac{1-(x-1)^k}{2-x}$.

The steady-state sectoral employment share of type ω workers are given by

$$\begin{pmatrix} \Pr(\text{sector } 1|\text{type } \omega) \\ \Pr(\text{sector } 2|\text{type } \omega) \end{pmatrix} = \begin{pmatrix} \frac{\beta_\omega}{\alpha_\omega + \beta_\omega} \\ \frac{\alpha_\omega}{\alpha_\omega + \beta_\omega} \end{pmatrix} \equiv \begin{pmatrix} \tilde{\beta}_\omega \\ \tilde{\alpha}_\omega \end{pmatrix},$$

which gives

$$\tilde{\ell}_1^\omega = \frac{\tilde{\beta}_\omega \theta_\omega}{\sum_{\omega'} \tilde{\beta}_{\omega'} \theta_{\omega'}} \quad \text{and} \quad \tilde{\ell}_2^\omega = \frac{\tilde{\alpha}_\omega \theta_\omega}{\sum_{\omega'} \tilde{\alpha}_{\omega'} \theta_{\omega'}}.$$

Thus, the k-period worker flow matrix is given by

$$\mathcal{F}_k = \begin{pmatrix} 1 - \dfrac{\sum_\omega \tilde{\beta}_\omega \theta_\omega \alpha_\omega f^k(\bar{\alpha}_\omega + \bar{\beta}_\omega)}{\sum_\omega \tilde{\beta}_\omega \theta_\omega} & \dfrac{\sum_\omega \tilde{\beta}_\omega \theta_\omega \alpha_\omega f^k(\bar{\alpha}_\omega + \bar{\beta}_\omega)}{\sum_\omega \tilde{\beta}_\omega \theta_\omega} \\ \dfrac{\sum_\omega \tilde{\alpha}_\omega \theta_\omega \beta_\omega f^k(\bar{\alpha}_\omega + \bar{\beta}_\omega)}{\sum_\omega \tilde{\alpha}_\omega \theta_\omega} & 1 - \dfrac{\sum_\omega \tilde{\alpha}_\omega \theta_\omega \beta_\omega f^k(\bar{\alpha}_\omega + \bar{\beta}_\omega)}{\sum_\omega \tilde{\alpha}_\omega \theta_\omega} \end{pmatrix}.$$

OA.2 Structural Estimation

OA.2.1 Extrapolation Using the Structural Model

We estimate the structural model by matching the observed worker flow matrices. 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. Following Assumption 2, we impose symmetry on the switching costs. The estimation process involves two steps: We first maximize the likelihood of observing $\{\mathcal{F}_k\}_{k=1}^{18}$ 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.

Table OA.1: Estimation Results

$\lvert\Omega\rvert = 2$	Type ω_1 (30.7%)					Type ω_2 (69.3%)				
Sector	Wage	Switching Cost				Wage	Switching Cost			
Agri/Const	0.58	0.00	1.54	1.69	1.55	1.02	0.00	4.62	5.63	5.46
Manufacturing	0.60	1.54	0.00	1.34	1.41	1.02	4.62	0.00	4.87	4.94
Commu/Trade	0.66	1.69	1.34	0.00	0.98	1.00	5.63	4.87	0.00	3.72
Services/Others	0.77	1.55	1.41	0.98	0.00	1.06	5.46	4.94	3.72	0.00

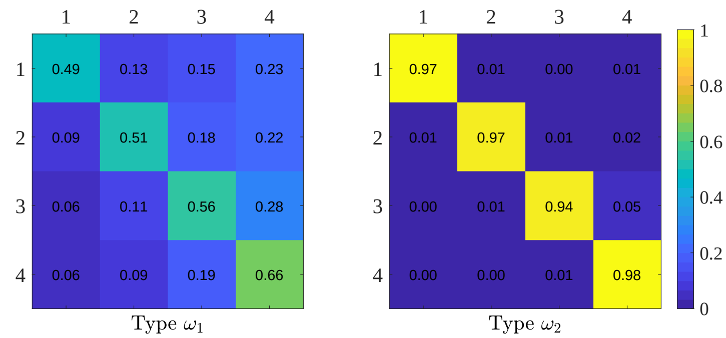


Figure OA.3. Type-Specific Transition Matrix

Table OA.1 shows the estimation result. The Bayesian information criterion supports the model with two worker types. Figure OA.3 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 OA.3. In contrast, the second type has much higher switching costs, so workers of this type rarely move to other sectors. In Section OA.2.3, we show how to interpret the figures in Table OA.1. In particular, paying one unit of switching costs means paying 3.25% of lifetime consumption. Thus, the switching costs in Table OA.1 are at most less than 20% of lifetime consumption. This is smaller than the estimates of Artuç, Chaudhuri, and McLaren (2010) (hereafter, 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.[2]

We also estimate primitives of the canonical model by matching the one-period worker flow matrix, $\mathcal{F}_1$; See Figure OA.4 for the results. All parameters, including elements of the transition matrix, lie between the corresponding parameters for the model with two worker types.

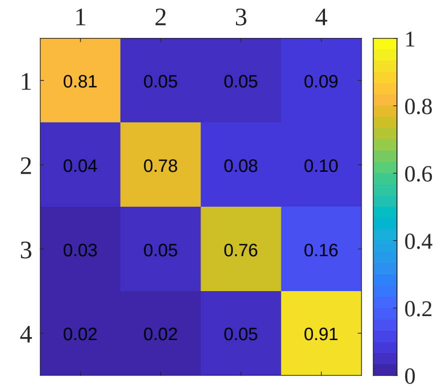
Model Fit. The fit of the model with two worker types, documented in Figures 3 and 4, is surprising for two reasons. First, the degrees of freedom (19 parameters) are much smaller than the number of moments we target (216 moments).[3] Second, the dynamic discrete choice framework imposes systematic restrictions on

[2] 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.

[3] 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

Wage	Switching Cost			
0.87	0.00	2.86	2.96	3.16
0.85	2.86	0.00	2.48	2.89
0.84	2.96	2.48	0.00	2.23
1.00	3.16	2.89	2.23	0.00

(a) Primitives



(b) Transition Matrix

Figure OA.4. Estimation Result: Canonical Model

Notes: Panel (a) shows the estimated values of the primitives of the canonical model. The four sectors are Agriculture and Construction; Manufacturing; Communications and Trade; and Services and Others. Panel (b) shows the resulting transition matrix (or, equivalently, one-year worker flow matrix).

the model-implied worker flow matrices, so we would not be able to match every worker flow matrix series even with an infinite number of worker types.

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 OA.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.

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

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.

(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.

OA.2.2 Alternative Methods of Extrapolation

Pure Extrapolation. Suppose we have panel data of length $K < \infty$. From the data, we can observe the conditional probability

$$\Pr\left(s_{t+k} = s_k | s_t = s_0, s_{t+1} = s_1, \cdots, s_{t+k-1} = s_{k-1}\right),$$

for all $s_0, s_1, \ldots, s_k \in \mathcal{S}$ and k less than K. To extrapolate the probabilities with k greater than or equal to K, we truncate the history and assume the following:

$$\begin{aligned}
&\Pr\left(s_{t+k} = s_k | s_t = s_0, s_{t+1} = s_1, \cdots, s_{t+k-1} = s_{k-1}\right) \\
&= \Pr\left(s_{t+k} = s_k | s_{t+k-K+1} = s_{k-K+1}, \cdots, s_{t+k-1} = s_{k-1}\right).
\end{aligned}$$

In short, we assume a $(K-1)$-th order Markov process and calculate the probabilities accordingly. In this sense, this extrapolation is a strict generalization of the canonical model's extrapolation, which is based on the assumption that the sector choice follow a first-order Markov process.

Extrapolation Using Retention Model. The retention model developed in Henry (1971) is based on the idea that Markov chains can be viewed as the conjunction of two processes: One determines whether workers change sectors or not, and the other governs which sector they choose, conditional on sector switching. Any transition matrix F can be decomposed as follows:

$$F = F^{\text{diag}} + (I - F^{\text{diag}})F^{\text{off-diag}},$$

where

$$\begin{aligned}
(F^{\text{diag}})_{i,j} &= F_{i,j} \cdot \mathbb{1}_{i=j}, \\
(F^{\text{off-diag}})_{i,j} &= \Pr(s_{t+1} = j | s_t = i, s_{t+1} \neq i) = \frac{F_{i,j}}{1 - F_{i,i}}.
\end{aligned}$$

The idea of extrapolation is to treat the two parts of the process separately. For example, we can extrapolate each element of the first part of the worker flow matrices to compute the full series of $\{(\mathcal{F}_k)^{\text{diag}}\}$. Then, we can assume that the second part remains constant for all k: $(\mathcal{F}_k)^{\text{off-diag}} = (\mathcal{F}_1)^{\text{off-diag}}$ for all k.

OA.2.3 Interpretation of the Estimation Results

Denote the log wages and switching costs estimated under the normalization $\rho = 1$ by $\ln w^0$ and C^0. Then, the true log wages and switching costs are given by

$$\ln w = \rho \cdot \ln w^0 \text{ and } C = \rho \cdot C^0.$$

Thus, when $C^0 = 1$, the consumption equivalent variation of paying switching cost is implicitly given by

$$\frac{\ln w(1 - \texttt{CEV})}{1 - \beta} = \frac{\ln w}{1 - \beta} - \rho,$$

which gives

$$\texttt{CEV} = 1 - \exp(-\rho(1 - \beta)) = 3.25\%.$$

Likewise, our estimate of Fréchet parameter is 0.825, which means that the consumption equivalent variation corresponding to one standard deviation lower realization of the idiosyncratic shock is given by

$$\frac{\ln w(1 - \texttt{CEV})}{1 - \beta} = \frac{\ln w}{1 - \beta} - 2 \cdot \frac{\pi\rho}{\sqrt{6}}$$

where we multiply two because it is the difference between two realizations of idiosyncratic shocks. This gives

$$\texttt{CEV} = 1 - \exp(-\frac{2\pi\rho(1 - \beta)}{\sqrt{6}}) = 8.12\%.$$

In contrast, ACM assume linear utility function, so the consumption equivalent variation can be computed as

$$\frac{w(1 - \texttt{CEV}_{\text{ACM}})}{1 - \beta} = \frac{w}{1 - \beta} - C_{\text{ACM}}$$

Thus, we have (given their normalization of average wages to one)

$$\texttt{CEV}_{\text{ACM}} = \frac{(1 - \beta)C_{\text{ACM}}}{w} = 19.7\%$$

where we use the number in Panel IV of Table 3, which is used in their counterfactual exercises. Likewise, their estimate of Fréchet parameter is 1.884, which means that the consumption equivalent variation corresponding to one standard deviation lower realization of the idiosyncratic shock is given by

$$\frac{w(1 - \texttt{CEV}_{\text{ACM}})}{1 - \beta} = \frac{w}{1 - \beta} - 2 \cdot \frac{\pi\rho_{\text{ACM}}}{\sqrt{6}}$$

which gives

$$\texttt{CEV}_{\text{ACM}} = \frac{2\pi\rho_{\text{ACM}}(1 - \beta)}{w\sqrt{6}} = 14.5\%.$$

OA.2.4 Estimation of ρ

Section 4.5 proposes a method to estimate the parameter ρ, which is based on the second equation of Proposition 1:

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

We refer to this relation as a forward-looking infinite-order MA process, because it resembles infinite-order MA processes, but involving forward-looking variables.

We first use equations (2) and (3) to prove equation (16) under the homogeneous worker assumption. Imposing Assumption 1 and assumptions in Lemma A.2, these equations become

$$\begin{aligned}\mathrm{d}v_t^{\omega} &= \mathrm{d}w_t + \beta F^{\omega}\,\mathbb{E}_t\,\mathrm{d}v_{t+1}^{\omega},\\ \mathrm{d}\ln\ell_{t+1}^{\omega} &= F^{\omega}\,\mathrm{d}\ln\ell_t^{\omega} + \tfrac{\beta}{\rho}(I - (F^{\omega})^2)\,\mathbb{E}_t\,\mathrm{d}v_{t+1}^{\omega}.\end{aligned} \tag{OA.2}$$

Pre-multiplying one-period forwarded version of equation (OA.2) with βF^{ω} and taking expectation operator $\mathbb{E}_t$, we have

$$\beta F^{\omega}\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2}^{\omega} = \beta(F^{\omega})^2\,\mathrm{d}\ln\ell_{t+1}^{\omega} + \tfrac{\beta}{\rho}(I-(F^{\omega})^2)\beta F^{\omega}\,\mathbb{E}_t\,\mathrm{d}v_{t+2}^{\omega}. \tag{OA.3}$$

Subtracting equation (OA.3) from equation (OA.2), we have

$$\mathrm{d}\ln\ell_{t+1}^{\omega} = F^{\omega}\,\mathrm{d}\ln\ell_t^{\omega} + \beta F^{\omega}(\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2}^{\omega} - F^{\omega}\,\mathrm{d}\ln\ell_{t+1}^{\omega}) + \tfrac{\beta}{\rho}(I-(F^{\omega})^2)\,\mathbb{E}_t\,\mathrm{d}w_{t+1}.$$

Under the homogeneous worker assumption, we can rearrange this result to obtain equation (16):

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

We refer to this equation as a *forward-looking process with order 2* because we write a period t variable as a function of period $t+1$ and $t+2$ variables and a shock.

Now, suppose that there are N number of worker types, $\omega \in \{1, 2, \ldots, N\}$. We derive a recursive representation of the form

$$\mathrm{d}\ln\ell_t = \sum_{k=1}^{K} \mathbf{\Gamma}_k\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+k} + \tfrac{\beta}{\rho}\sum_{k=1}^{K'+1}\mathbf{\Lambda}_k\,\mathbb{E}_t\,\mathrm{d}w_{t+k}.$$

We refer to this equation as a forward-looking process with order (K, K'). Note that the change in aggregate labor supply can be written as

$$\mathrm{d}\ln\ell_{t+1} = \bar{\mathbb{E}}_{\omega}[\mathrm{d}\ln\ell_{t+1}^{\omega}] = L_1\,\mathrm{d}\ln\ell_{t+1}^{1} + L_2\,\mathrm{d}\ln\ell_{t+1}^{2} + \cdots + L_N\,\mathrm{d}\ln\ell_{t+1}^{N},$$

where each $\mathrm{d}\ln\ell_{t+1}^{\omega}$ follows a forward-looking process with order w, and L_{ω} is a diagonal matrix, whose i-th diagonal element is given by the steady-state proportion of type ω in sector i. Granger and Morris (1976) show that the scalar-weighted sum of N number of autoregressive processes of order 2 follows an autoregressive process of order at most $2N$. Here, we instead have diagonal-matrix-weighted sum of N number of forward-looking process of order 2, but we can apply a modified version of their proof to show proposition. Due to an invertibility issue, we need forward-looking process of order $(4N-2, 4N-4)$ instead of order $2N$. The next subsection is devoted to the proof of Proposition OA.3.

Proposition OA.3. *If the number of types is* N*,* $\mathrm{d}\ln\ell_t$ *has a recursive representation of the form*

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

OA.2.4.1 Proof of Proposition OA.3

We prove this proposition for the case with $N = 2$. The same proof can be inductively applied to show the case with $N > 2$. For the case with two worker types, aggregate labor supply is given by

$$\mathrm{d}\ln\ell_t = L_1 \,\mathrm{d}\ln\ell_t^1 + L_2 \,\mathrm{d}\ln\ell_t^2, \tag{OA.4}$$

where

$$\mathrm{d}\ln\ell_{t+1}^1 - \beta F^1 \,\mathbb{E}_t \,\mathrm{d}\ln\ell_{t+2}^1 = F^1 \,\mathrm{d}\ln\ell_t^1 - \beta(F^1)^2 \,\mathrm{d}\ln\ell_{t+1}^1 + \tfrac{\beta}{\rho}(I-(F^1)^2)\,\mathbb{E}_t \,\mathrm{d}w_{t+1}, \tag{OA.5}$$

$$\mathrm{d}\ln\ell_{t+1}^2 - \beta F^2 \,\mathbb{E}_t \,\mathrm{d}\ln\ell_{t+2}^2 = F^2 \,\mathrm{d}\ln\ell_t^2 - \beta(F^2)^2 \,\mathrm{d}\ln\ell_{t+1}^2 + \tfrac{\beta}{\rho}(I-(F^2)^2)\,\mathbb{E}_t \,\mathrm{d}w_{t+1}. \tag{OA.6}$$

Using equation (OA.5) to cancel out $\mathrm{d}\ln\ell_t^1$ from equation (OA.4), we have

$$\begin{aligned}
&L_1((F^1)^{-1} + \beta F^1)L_1^{-1}\,\mathrm{d}\ln\ell_{t+1} - \beta\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2} - \mathrm{d}\ln\ell_t \\
&= L_1((F^1)^{-1} + \beta F^1)(\mathrm{d}\ln\ell_{t+1}^1 + L_1^{-1}L_2\,\mathrm{d}\ln\ell_{t+1}^2) - \beta\,\mathbb{E}_t(L_1\,\mathrm{d}\ln\ell_{t+2}^1 + L_2\,\mathrm{d}\ln\ell_{t+2}^2) - (L_1\,\mathrm{d}\ln\ell_t^1 + L_2\,\mathrm{d}\ln\ell_t^2) \\
&= L_1(F^1)^{-1}\big((I+\beta(F^1)^2)\,\mathrm{d}\ln\ell_{t+1}^1 - \beta F^1\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2}^1 - F^1\,\mathrm{d}\ln\ell_t^1\big) + \mathbb{E}_t\,\Xi_t \\
&= L_1(F^1)^{-1}\tfrac{\beta}{\rho}(I-(F^1)^2)\,\mathbb{E}_t\,\mathrm{d}w_{t+1} + \mathbb{E}_t\,\Xi_t,
\end{aligned}$$

where

$$\begin{aligned}
\Xi_t &= L_1((F^1)^{-1} + \beta F^1)L_1^{-1}L_2\,\mathrm{d}\ln\ell_{t+1}^2 - \beta L_2\,\mathrm{d}\ln\ell_{t+2}^2 - L_2\,\mathrm{d}\ln\ell_t^2 \\
&= L_1((F^1)^{-1} + \beta F^1)L_1^{-1}L_2\,\mathrm{d}\ln\ell_{t+1}^2 - \beta L_2\,\mathrm{d}\ln\ell_{t+2}^2 \\
&\quad - L_2(F^2)^{-1}\Big(((I+\beta(F^2)^2)\,\mathrm{d}\ln\ell_{t+1}^2 - \beta F^2\,\mathrm{d}\ln\ell_{t+2}^2 - \tfrac{\beta}{\rho}(I-(F^2)^2)\,\mathbb{E}_t\,\mathrm{d}w_{t+1}\Big) \\
&= \Big(L_1((F^1)^{-1} + \beta F^1)L_1^{-1}L_2 - L_2((F^2)^{-1} + \beta F^2)\Big)\,\mathrm{d}\ln\ell_{t+1}^2 + \tfrac{\beta}{\rho}L_2((F^2)^{-1} - F^2)\,\mathbb{E}_t\,\mathrm{d}w_{t+1}.
\end{aligned}$$

This can be rearranged to

$$\begin{aligned}
&L_1((F^1)^{-1} + \beta F^1)L_1^{-1}\,\mathrm{d}\ln\ell_{t+1} - \beta\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2} - \mathrm{d}\ln\ell_t \\
&= \tfrac{\beta}{\rho}\big(L_1((F^1)^{-1} - F^1) + L_2((F^2)^{-1} - F^2)\big)\,\mathbb{E}_t\,\mathrm{d}w_{t+1} + \big(L_1((F^1)^{-1} + \beta F^1)L_1^{-1} - L_2((F^2)^{-1} + \beta F^2)L_2^{-1}\big)L_2\,\mathrm{d}\ln\ell_{t+1}^2,
\end{aligned}$$

or equivalently $y_t = \Psi x_t$ where

$$y_t = \boldsymbol{X}\,\mathrm{d}\ln\ell_{t+1} - \beta\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2} - \mathrm{d}\ln\ell_t - \tfrac{\beta}{\rho}\boldsymbol{Y}\,\mathbb{E}_t\,\mathrm{d}w_{t+1},$$

$$
\begin{aligned}
x_t &= \mathrm{d}\ln \ell_{t+1}^2, \\
\boldsymbol{X} &= L_1((F^1)^{-1} + \beta F^1)L_1^{-1}, \\
\boldsymbol{Y} &= L_1((F^1)^{-1} - F^1) + L_2((F^2)^{-1} - F^2), \\
\Psi &= \big(L_1((F^1)^{-1} + \beta F^1)L_1^{-1} - L_2((F^2)^{-1} + \beta F^2)L_2^{-1}\big)L_2.
\end{aligned}
$$

From equation (OA.6), the law of motion of x_t is given by

$$
x_t = \boldsymbol{A} x_{t+1} + \boldsymbol{B}\,\mathbb{E}_{t+1}\, x_{t+2} + \varepsilon_{t+1},
$$

where $\boldsymbol{A} = (F^2)^{-1} + \beta F^2$, $\boldsymbol{B} = -\beta I$, and $\varepsilon_t = \frac{\beta}{\rho}\boldsymbol{C}\,\mathbb{E}_t\,\mathrm{d}w_{t+1}$ where $\boldsymbol{C} = -((F^2)^{-1} - F^2)$.

Lemma OA.2. *Suppose $x_t = \boldsymbol{A} x_{t+1} + \boldsymbol{B} x_{t+2} + \varepsilon_{t+1} \in \mathbb{R}^S$ and $y_t = \boldsymbol{Z} x_t$ where $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{S\times S}$ are invertible and $\boldsymbol{Z} \in \mathbb{R}^{S\times S}$ is of rank $S-1$. Then, we can always write*

$$
y_t = \Theta_1 y_{t+1} + \Theta_2 y_{t+2} + \Theta_3 y_{t+3} + \Theta_4 y_{t+4} + \Omega_1 \varepsilon_{t+1} + \Omega_2 \varepsilon_{t+2} + \Omega_3 \varepsilon_{t+3}
$$

for some matrices $\Theta_i = \Theta_i(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{Z})$ and $\Omega_i = \Omega_i(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{Z})$.

Proof. Write $\boldsymbol{x} = \begin{pmatrix} x_{t+3} \\ x_{t+4} \end{pmatrix}$, then we have

$$
\begin{aligned}
y_{t+4} &= \begin{pmatrix} \boldsymbol{O} & \boldsymbol{Z} \end{pmatrix} \boldsymbol{x} \equiv \boldsymbol{M}_1 \boldsymbol{x}, \\
y_{t+3} &= \begin{pmatrix} \boldsymbol{Z} & \boldsymbol{O} \end{pmatrix} \boldsymbol{x} \equiv \boldsymbol{M}_2 \boldsymbol{x}, \\
y_{t+2} &= \boldsymbol{Z}(\boldsymbol{A} x_{t+3} + \boldsymbol{B} x_{t+4} + \varepsilon_{t+3}) = \begin{pmatrix} \boldsymbol{ZA} & \boldsymbol{ZB} \end{pmatrix} \boldsymbol{x} + \boldsymbol{Z}\varepsilon_{t+3} \equiv \boldsymbol{M}_3 \boldsymbol{x} + \boldsymbol{Z}\varepsilon_{t+3}, \\
y_{t+1} &= \boldsymbol{Z}(\boldsymbol{A} x_{t+2} + \boldsymbol{B} x_{t+3} + \varepsilon_{t+2}) = \boldsymbol{Z}(\boldsymbol{A}(\boldsymbol{A} x_{t+3} + \boldsymbol{B} x_{t+4} + \varepsilon_{t+3}) + \boldsymbol{B} x_{t+3} + \varepsilon_{t+2}) \\
&= \begin{pmatrix} \boldsymbol{Z}(\boldsymbol{A}^2 + \boldsymbol{B}) & \boldsymbol{ZAB} \end{pmatrix} \boldsymbol{x} + \boldsymbol{ZA}\varepsilon_{t+3} + \boldsymbol{Z}\varepsilon_{t+2} \equiv \boldsymbol{M}_4 \boldsymbol{x} + \boldsymbol{ZA}\varepsilon_{t+3} + \boldsymbol{Z}\varepsilon_{t+2}, \\
y_t &= \boldsymbol{Z}(\boldsymbol{A} x_{t+1} + \boldsymbol{B} x_{t+2} + \varepsilon_{t+1}) \\
&= \boldsymbol{Z}(\boldsymbol{A}(\boldsymbol{A}(\boldsymbol{A} x_{t+3} + \boldsymbol{B} x_{t+4} + \varepsilon_{t+3}) + \boldsymbol{B} x_{t+3} + \varepsilon_{t+2}) + \boldsymbol{B}(\boldsymbol{A} x_{t+3} + \boldsymbol{B} x_{t+4} + \varepsilon_{t+3}) + \varepsilon_{t+1}) \\
&= \begin{pmatrix} \boldsymbol{Z}(\boldsymbol{A}^3 + \boldsymbol{AB} + \boldsymbol{BA}) & \boldsymbol{Z}(\boldsymbol{A}^2\boldsymbol{B} + \boldsymbol{B}^2) \end{pmatrix} \boldsymbol{x} + \boldsymbol{Z}(\boldsymbol{A}^2 + \boldsymbol{B})\varepsilon_{t+3} + \boldsymbol{ZA}\varepsilon_{t+2} + \boldsymbol{Z}\varepsilon_{t+1} \\
&\equiv \boldsymbol{M}_5 \boldsymbol{x} + \boldsymbol{Z}(\boldsymbol{A}^2 + \boldsymbol{B})\varepsilon_{t+3} + \boldsymbol{ZA}\varepsilon_{t+2} + \boldsymbol{Z}\varepsilon_{t+1}.
\end{aligned}
$$

Note that $\boldsymbol{N}_1 \equiv \boldsymbol{M}_1$ is of rank $S-1$, $\boldsymbol{N}_2 \equiv \begin{pmatrix} \boldsymbol{M}_1 \\ \boldsymbol{M}_2 \end{pmatrix}$ is of rank $2(S-1)$. If all rows of $\boldsymbol{M}_3$ can be written as a linear combination of rows of $\boldsymbol{N}_2$, then we can write $y_t = \Theta_1 y_{t+1} + \Theta_2 y_{t+2}$. If not, $\boldsymbol{N}_3 \equiv \begin{pmatrix} \boldsymbol{N}_2 \\ \boldsymbol{M}_3 \end{pmatrix}$ is of rank at least $2(S-1)+1$. By the same logic, we either prove the result, or $\boldsymbol{N}_4 \equiv \begin{pmatrix} \boldsymbol{N}_3 \\ \boldsymbol{M}_4 \end{pmatrix}$ is of rank (at least) $2S$. Thus, all rows of $\boldsymbol{M}_5$ can be written as a linear combination of rows of $\boldsymbol{N}_4$. This implies that we can find matrices $\Theta_1, \cdots, \Theta_4$ such that

$$
\boldsymbol{M}_5 = \Theta_1 \boldsymbol{M}_4 + \Theta_2 \boldsymbol{M}_3 + \Theta_3 \boldsymbol{M}_2 + \Theta_4 \boldsymbol{M}_1.
$$

Thus, we only need

$$\boldsymbol{Z}(\boldsymbol{A}^2+\boldsymbol{B})\varepsilon_{t+3}+\boldsymbol{Z}\boldsymbol{A}\varepsilon_{t+2}+\boldsymbol{Z}\varepsilon_{t+1}=\Theta_1(\boldsymbol{Z}\boldsymbol{A}\varepsilon_{t+3}+\boldsymbol{Z}\varepsilon_{t+2})+\Theta_2(\boldsymbol{Z}\varepsilon_{t+3})+\Omega_1\varepsilon_{t+1}+\Omega_2\varepsilon_{t+2}+\Omega_3\varepsilon_{t+3}$$

or

$$\Omega_1=\boldsymbol{Z},\ \ \Omega_2=\boldsymbol{Z}\boldsymbol{A}-\Theta_1\boldsymbol{Z},\ \text{ and }\ \Omega_3=\boldsymbol{Z}(\boldsymbol{A}^2+\boldsymbol{B})-\Theta_1\boldsymbol{Z}\boldsymbol{A}-\Theta_2\boldsymbol{Z}. \qquad \square$$

Lemma OA.3. Ψ *is non-invertible.*

Proof. Denote population share of type ω as θ^ω, transition matrix as F^ω, and stationary distribution over sectors as π^ω. Define $\pi\equiv\sum_\omega\theta^\omega\pi^\omega$, then we have

$$\pi^\intercal L_\omega=\theta^\omega(\pi^\omega)^\intercal\ \text{ and }\ (\pi^\omega)^\intercal F^\omega=(\pi^\omega)^\intercal.$$

Thus,

$$\pi^\intercal\Psi=\pi^\intercal\Big(L_1((F^1)^{-1}+\beta F^1)L_1^{-1}-L_2((F^2)^{-1}+\beta F^2)L_2^{-1}\Big)L_2=(\pi^\intercal+\beta\pi^\intercal-\pi^\intercal-\beta\pi^\intercal)L_2=0. \quad \square$$

By combining the previous two lemmas, we can write

$$y_t=\Theta_1y_{t+1}+\Theta_2\,\mathbb{E}_{t+1}\,y_{t+2}+\Theta_3\,\mathbb{E}_{t+1}\,y_{t+3}+\Theta_4\,\mathbb{E}_{t+1}\,y_{t+4}+\Omega_1\varepsilon_{t+1}+\Omega_2\,\mathbb{E}_{t+1}\,\varepsilon_{t+2}+\Omega_3\,\mathbb{E}_{t+1}\,\varepsilon_{t+3},$$

where $\Theta_i=\Theta_i(\boldsymbol{A},\boldsymbol{B},\Psi)$ and $\Omega_i=\Omega_i(\boldsymbol{A},\boldsymbol{B},\Psi)$. Plugging in the definitions of y_t and ε_t, we have

$$\begin{aligned}-\mathrm{d}\ln\ell_t+\boldsymbol{X}\mathrm{d}\ln\ell_{t+1}-\beta\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2}-\tfrac{\beta}{\rho}\boldsymbol{Y}\mathbb{E}_t\,\mathrm{d}w_{t+1}=\;&\Theta_1(-\mathrm{d}\ln\ell_{t+1}+\boldsymbol{X}\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2}-\beta\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+3}-\tfrac{\beta}{\rho}\boldsymbol{Y}\mathbb{E}_t\,\mathrm{d}w_{t+2})\\&+\Theta_2(-\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2}+\boldsymbol{X}\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+3}-\beta\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+4}-\tfrac{\beta}{\rho}\boldsymbol{Y}\mathbb{E}_t\,\mathrm{d}w_{t+3})\\&+\Theta_3(-\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+3}+\boldsymbol{X}\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+4}-\beta\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+5}-\tfrac{\beta}{\rho}\boldsymbol{Y}\mathbb{E}_t\,\mathrm{d}w_{t+4})\\&+\Theta_4(-\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+4}+\boldsymbol{X}\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+5}-\beta\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+6}-\tfrac{\beta}{\rho}\boldsymbol{Y}\mathbb{E}_t\,\mathrm{d}w_{t+5})\\&+\tfrac{\beta}{\rho}\Omega_1\boldsymbol{C}\mathbb{E}_t\,\mathrm{d}w_{t+2}+\tfrac{\beta}{\rho}\Omega_2\boldsymbol{C}\mathbb{E}_t\,\mathrm{d}w_{t+3}+\tfrac{\beta}{\rho}\Omega_3\boldsymbol{C}\mathbb{E}_t\,\mathrm{d}w_{t+4}\end{aligned}$$

or equivalently

$$\begin{aligned}\mathrm{d}\ln\ell_t=\;&(\Theta_1+\boldsymbol{X})\,\mathrm{d}\ln\ell_{t+1}+(-\Theta_1\boldsymbol{X}+\Theta_2-\beta I)\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2}+(\beta\Theta_1-\Theta_2\boldsymbol{X}+\Theta_3)\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+3}\\&+(\beta\Theta_2-\Theta_3\boldsymbol{X}+\Theta_4)\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+4}+(\beta\Theta_3-\Theta_4\boldsymbol{X})\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+5}+(\beta\Theta_4)\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+6}\\&-\tfrac{\beta}{\rho}\boldsymbol{Y}\mathbb{E}_t\,\mathrm{d}w_{t+1}+(\tfrac{\beta}{\rho}\Theta_1\boldsymbol{Y}-\tfrac{\beta}{\rho}\Omega_1\boldsymbol{C})\,\mathbb{E}_t\,\mathrm{d}w_{t+2}+(\tfrac{\beta}{\rho}\Theta_2\boldsymbol{Y}-\tfrac{\beta}{\rho}\Omega_2\boldsymbol{C})\,\mathbb{E}_t\,\mathrm{d}w_{t+3}+(\tfrac{\beta}{\rho}\Theta_3\boldsymbol{Y}-\tfrac{\beta}{\rho}\Omega_3\boldsymbol{C})\,\mathbb{E}_t\,\mathrm{d}w_{t+4}+\tfrac{\beta}{\rho}\Theta_4\boldsymbol{Y}\mathbb{E}_t\,\mathrm{d}w_{t+5}.\end{aligned}$$

This can be rearranged to obatin Proposition OA.3.[4]

$$\begin{aligned}\mathrm{d}\ln\ell_t=\;&\boldsymbol{\Gamma}_1\,\mathrm{d}\ln\ell_{t+1}+\boldsymbol{\Gamma}_2\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+2}+\boldsymbol{\Gamma}_3\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+3}+\boldsymbol{\Gamma}_4\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+4}+\boldsymbol{\Gamma}_5\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+5}+\boldsymbol{\Gamma}_6\,\mathbb{E}_t\,\mathrm{d}\ln\ell_{t+6}\\&+\tfrac{\beta}{\rho}(\boldsymbol{\Lambda}_1\,\mathbb{E}_t\,\mathrm{d}w_{t+1}+\boldsymbol{\Lambda}_2\,\mathbb{E}_t\,\mathrm{d}w_{t+2}+\boldsymbol{\Lambda}_3\,\mathbb{E}_t\,\mathrm{d}w_{t+3}+\boldsymbol{\Lambda}_4\,\mathbb{E}_t\,\mathrm{d}w_{t+4}+\boldsymbol{\Lambda}_5\,\mathbb{E}_t\,\mathrm{d}w_{t+5}).\end{aligned}$$

[4] If the number of types is 3, then we need $(\mathrm{d}\ln\ell_{t+1},\cdots,\mathrm{d}\ln\ell_{t+10})$, and with the number of types W, we need $(\mathrm{d}\ln\ell_{t+1},\cdots,\mathrm{d}\ln\ell_{t+4W-2})$.

OA.3 Applications

OA.3.1 Application 1: Hypothetical Trade Liberalization

Closing the Model. We consider the first version of the model in ACM, in which all sectors produce tradable output, and world prices are exogenously determined. Each sector s has a constant elasticity of substitution (CES) production function

$$y_t^s = \psi^s(\alpha^s(L_t^s)^{\rho^s} + (1-\alpha^s)(K_t^s)^{\rho^s})^{\frac{1}{\rho^s}}$$

with fixed sector-specific capital $K_t^s = 1$ normalized to one. The parameters should satisfy $\alpha^s \in (0,1)$, $\rho^s < 1$, and $\psi^s > 0$. Then, the sectoral wage is given by the marginal productivity of labor

$$w_t^s = p_t^s \alpha^s \psi^s (L_t^s)^{\rho^s - 1}(\alpha^s(L_t^s)^{\rho^s} + (1-\alpha^s))^{\frac{1-\rho^s}{\rho^s}},$$

where p_t^s is the domestic price of the sector-s good. Without loss of generality, we normalize the domestic prices to one in the initial steady state prior to the shock. Finally, workers have identical Cobb-Douglas with shares μ^s for sector s.

We follow the calibration strategy of ACM (except for the number of sectors). We set the values of α^s, ρ^s, and ψ^s to minimize the Euclidean distance between the model-implied values of sectoral wages, sectoral labor shares, and sector share of GDP; and the values computed from the data. The values of μ^s are calibrated using consumption shares from the Bureau of Labor Statistics (BLS). The obtained parameter values are summarized in Table OA.2.

Table OA.2: Parameter Values

Sector	α^s	ρ^s	ψ^s	μ^s
Agri./Const.	0.637	0.517	0.684	0.37
Manufacturing	0.420	0.487	1.094	0.3
Commu./Trade	0.600	0.561	1.069	0.08
Services/Others	0.401	0.530	1.442	0.25

OA.3.2 Application 2: The China Shock

OA.3.2.1 CDP's Model Extended with Worker Heterogeneity

In this section and the next, we closely follow the modeling decisions and notation of Caliendo, Dvorkin, and Parro (2019) (hereafter, CDP). See CDP for more details and equilibrium characterization. CDP consider a world with N locations (indexed by n or i) and J sectors (indexed by j or k), where sector $j = 0$ represents non-employment. Time is discrete and indexed by $t \in \mathbb{N}_0$. In each location-sector combination, (n, j), there is a competitive local labor market.

Heterogeneous Workers. In each location n, there is a continuum of forward-looking workers who optimally decide which sector to supply their labor for each period. The heterogeneity of workers are indexed by ω. Similar to equation (1), the value of a worker of type ω who is employed in sector j at period t is given by

$$V_{jt}^{n,\omega} = U(c_{jt}^{n,\omega}) + \max_k \{\beta \, \mathbb{E}_t \, V_{kt+1}^{n,\omega} - C_{jk}^{n,\omega} + \rho \cdot \varepsilon_{kt}\}$$

where $c_{jt}^{n,\omega} = \prod_k (c_{jt}^{k,n,\omega})^{\alpha^k}$ is a Cobb-Douglas aggregator across local sectoral goods, with the corresponding price index $P_t^n = \prod_k (P_t^{nk}/\alpha^k)^{\alpha^k}$. Following CDP, we assume $U(c) = \log c$.

Households employed in a local labor market (n, j) earn a nominal wage of w_t^{nj}, consuming $c_{jt}^{n,\omega} = \tilde{c}^{nj,\omega} \cdot w_t^{nj}/P_t^n$ units of consumption aggregate where $\tilde{c}^{nj,\omega}$ is a type-specific shifter representing non-pecuniary sectoral preferences. Non-employed household (who chooses sector $j = 0$) consumes $c_{0t}^{n,\omega} = b^{n,\omega}$ units of consumption aggregate in terms of home production. The ex-ante value and transition probabilities are characterized as in equations (2) and (3):

$$v_{jt}^{n,\omega} = U(c_{jt}^{n,\omega}) + \rho \ln \sum_k (\exp(\beta \, \mathbb{E}_t \, v_{kt+1}^{n,\omega}) / \exp(C_{jk}^{n,\omega}))^{1/\rho}, \tag{OA.7}$$

$$F_{jkt}^{n,\omega} = \frac{(\exp(\beta \, \mathbb{E}_t \, v_{kt+1}^{n,\omega}) / \exp(C_{jk}^{n,\omega}))^{1/\rho}}{\sum_{k'} (\exp(\beta \, \mathbb{E}_t \, v_{k't+1}^{n,\omega}) / \exp(C_{jk'}^{n,\omega}))^{1/\rho}}. \tag{OA.8}$$

Thus, the law of motion of sectoral labor supply of type ω workers in region n is

$$\ell_{kt+1}^{n,\omega} = \sum_j F_{jkt}^{n,\omega} \ell_{jt}^{n,\omega}. \tag{OA.9}$$

Finally, let $L_t^{nj} = \sum_\omega \ell_{jt+1}^{n,\omega}$ be the total labor supply to local labor market (n, j). In this environment, Assumption 1 holds, and we implicitly maintain Assumption 2 (with n superscript) as well.

Production. For each sector j, there is a continuum of different varieties of intermediate goods. Each region-sector combination draws a variety-specific productivity z^{nj}, which follows a Fréchet distribution with the dispersion parameter θ^j. Without loss, each variety is indexed by $z^j = (z^{1j}, z^{2j}, \ldots, z^{Nj})$. In each local labor market, (n, j), there is a continuum of perfectly competitive firms producing variety z^j. They have a Cobb-Douglas technology combining labor (l), structures (h), and local sectoral goods from all sectors (M):

$$q_t^{nj} = z^{nj} (A_t^{nj} (h_t^{nj})^{\xi^n} (l_t^{nj})^{1-\xi^n})^{\gamma^{nj}} \prod_k (M_t^{nj,nk})^{\gamma^{nj,nk}},$$

where A_t^{nj} is a sector-region specific productivity. Thus, the unit cost of producing this intermediate good is

$$\frac{x_t^{nj}}{z^{nj} (A_t^{nj})^{\gamma^{nj}}} \quad \text{where } x_t^{nj} = B^{nj} ((r_t^{nj})^{\xi^n} (w_t^{nj})^{1-\xi^n})^{\gamma^{nj}} \prod_k (P_t^{nk})^{\gamma^{nj,nk}}, \tag{OA.10}$$

where B^{nj} is a constant, r_t^{nj} is the rental price of structures, and P_t^{nk} is the price of the local sector-k goods.

Local sectoral goods are produced from intermediate goods in a competitive way, which are then used as final consumption and as materials for the production of intermediate varieties. The technology is given by:

$$Q_t^{nj} = \left(\int (\tilde{q}_t^{nj}(z^j))^{1-1/\eta^{nj}} \, \mathrm{d}\phi^j(z^j) \right)^{\eta^{nj}/(\eta^{nj}-1)},$$

where $\tilde{q}_t^{nj}(z^j)$ is the quantity of variety z^j used in the production, and $\phi^j(\cdot)$ is the joint distribution of the vector z^j. The intermediate good of variety z^j is sourced from a country with the minimum price, taking into account bilateral iceberg-type trade costs (κ). The minimized price is then given by

$$p_t^{nj}(z^j) = \min_i \left\{ \frac{\kappa_t^{nj,ij} x_t^{ij}}{z^{ij}(A_t^{ij})^{\gamma^{ij}}} \right\}.$$

Thus, the price of the local sectoral good is

$$P_t^{nj} = \Gamma\left(\frac{1+\theta^j-\eta^{nj}}{\theta^j}\right) \cdot \left(\sum_i (x_t^{ij}\kappa_t^{nj,ij})^{-\theta^j}(A_t^{ij})^{\theta^j\gamma^{ij}} \right)^{-1/\theta^j}. \tag{OA.11}$$

Finally, the share of total expenditure in local market (n, j) on goods from market (i, j) is given by

$$\pi_t^{nj,ij} = \frac{(x_t^{ij}\kappa_t^{nj,ij})^{-\theta^j}(A_t^{ij})^{\theta^j\gamma^{ij}}}{\sum_{i'}(x_t^{i'j}\kappa_t^{nj,i'j})^{-\theta^j}(A_t^{i'j})^{\theta^j\gamma^{i'j}}}. \tag{OA.12}$$

Structure Rentier. There is a continuum of structure rentiers in each region n. They own the local structures of fixed amount $\{H^{nj}\}_j$ and rent them to local firms. The received rents are aggregated at the global-level, and rentiers in each region n receive a constant share ι^n of the total global revenue:

$$\iota^n \chi_t \quad \text{where } \chi_t = \sum_i \sum_k r_t^{ik} H^{ik}.$$

Market Clearing. Market clearing for goods market, labor market, and structure market is given by

$$X_t^{nj} = \sum_k \gamma^{nk,nj} \sum_i \pi_t^{ik,nk} X_t^{ik} + \alpha^j \left(\sum_k w_t^{nk} L_t^{nk} + \iota^n \chi_t \right), \tag{OA.13}$$

$$w_t^{nj} L_t^{nj} = \gamma^{nj}(1-\xi^n) \sum_i \pi_t^{ij,nj} X_t^{ij}, \tag{OA.14}$$

$$r_t^{nj} H^{nj} = \gamma^{nj} \xi^n \sum_i \pi_t^{ij,nj} X_t^{ij}, \tag{OA.15}$$

where X_t^{nj} is the total expenditure on sector j good in region n.

Equilibrium. Following CDP, we group exogenous state variables of the economy into time-varying ones and time-invariant ones:

$$\Theta_t \equiv (\{A_t^{nj}\}_{n,j}, \{\kappa_t^{nj,ij}\}_{n,i,j}) \text{ and } \bar{\Theta} \equiv (\{C_{jk}^{n,\omega}\}_{j,k,n,\omega}, \{H^{nj}\}_{n,j}, \{\tilde{c}^{nj,\omega}\}_{n,j,\omega}, \{b^{n,\omega}\}_{n,\omega}).$$

Given the initial distribution of labor and the path of exogenous state variables $(\{\ell_{j0}^{n,\omega}\}_{j,n,\omega}, \{\Theta_t\}_{t=0}^{\infty}, \bar{\Theta})$, a *sequential competitive equilibrium* corresponds to a sequence of $\{L_t, F_t, v_t, x_t, P_t, \pi_t, X_t, w_t, r_t\}_{t=0}^{\infty}$, where $L_t = \{\ell_{jt}^{n,\omega}\}_{j,n,\omega}$, $F_t = \{F_{jkt}^{n,\omega}\}_{j,k,n,\omega}$, $v_t = \{v_{jt}^{n,\omega}\}_{j,n,\omega}$, $x_t = \{x_t^{nj}\}_{n,j}$, $P_t = \{P_t^{n,j}\}_{n,j}$, $\pi_t = \{\pi_t^{ij,nj}\}_{i,j,n}$, $X_t = \{X_t^{nj}\}_{n,j}$, $w_t = \{w_t^{nj}\}_{n,j}$, and $r_t = \{r_t^{nj}\}_{n,j}$, such that households optimally make sector choice decisions, as described in (OA.7)–(OA.9); firms maximize their profits, as described in (OA.10)–(OA.12); all markets clear, as described in (OA.13)–(OA.15). A *stationary equilibrium* is a sequential competitive equilibrium such that $\{L_t, F_t, v_t, x_t, P_t, \pi_t, X_t, w_t, r_t\}$ is time-invariant.

OA.3.2.2 Dynamic Hat Algebra with Worker Heterogeneity

Following CDP, we solve for the equilibrium in time differences. We denote by $\dot{y}_{t+1} \equiv (y_{t+1}^1/y_t^1, y_{t+1}^2/y_t^2, \cdots)$ the proportional change in any scalar or vector. The following proposition corresponds to Propositions 1 and 2 of CDP, but allowing for worker heterogeneity.

Proposition OA.4 (Solving the model). *Suppose that the economy is initially starting from a stationary equilibrium at period $t = 0$. Up to the first-order approximation around a stationary equilibrium, given a sequence of changes in exogenous state variables, $\{\dot{\Theta}_t\}_{t=1}^{\infty}$ satisfying $\lim_{t\to\infty} \dot{\Theta}_t = 1$, known to agents in period $t = 1$, the solution to the sequential equilibrium in time differences does not require information on the level of the exogenous state variables $\{\Theta_t\}_{t=0}^{\infty}$ or $\bar{\Theta}$, and solves the following system of equations:*

$$v_t^{nj} = v_0^{nj} + \sum_k \beta^k \mathcal{F}_k \ln\left(\frac{\dot{w}_{t+k}^n}{\dot{P}_{t+k}^n} \cdot \frac{\dot{w}_{t+k-1}^n}{\dot{P}_{t+k-1}^n} \cdot \dots \cdot \frac{\dot{w}_1^n}{\dot{P}_1^n}\right),$$

$$L_t^{nj} = L_0^{nj} \cdot \exp\left(\left(\sum_{s=0}^{t-2}\sum_{k=0}^{\infty} \frac{\beta^{k+1}}{\rho}(\mathcal{F}_{s+k}^n - \mathcal{F}_{s+k+2}^n) \ln\left(\frac{\dot{w}_{t-s+k}^n}{\dot{P}_{t-s+k}^n} \cdot \frac{\dot{w}_{t-s+k-1}^n}{\dot{P}_{t-s+k-1}^n} \cdot \dots \cdot \frac{\dot{w}_1^n}{\dot{P}_1^n}\right)\right)_j\right),$$

$$\dot{x}_{t+1}^{nj} = (\dot{L}_{t+1}^{nj})^{\gamma^{nj}\xi^n} (\dot{w}_{t+1}^{nj})^{\gamma^{nj}} \prod_k (\dot{P}_{t+1}^{nk})^{\gamma^{nj,nk}},$$

$$\dot{P}_{t+1}^{nj} = \left(\sum_i \pi_t^{nj,ij} (\dot{x}_{t+1}^{ij}\dot{\kappa}_{t+1}^{nj,ij})^{-\theta^j} (\dot{A}_{t+1}^{ij})^{\theta^j\gamma^{ij}}\right)^{-1/\theta^j},$$

$$\pi_{t+1}^{nj,ij} = \pi_t^{nj,ij} \left(\frac{\dot{x}_{t+1}^{ij}\dot{\kappa}_{t+1}^{nj,ij}}{\dot{P}_{t+1}^{nj}}\right)^{-\theta^j} (\dot{A}_{t+1}^{ij})^{\theta^j\gamma^{ij}},$$

$$X_{t+1}^{nj} = \sum_k \gamma^{nk,nj} \sum_i \pi_{t+1}^{ik,nk} X_{t+1}^{ik} + \alpha^j \left(\sum_k \dot{w}_{t+1}^{nk}\dot{L}_{t+1}^{nk} w_t^{nk} L_t^{nk} + \iota^n \chi_{t+1}\right),$$

$$\dot{w}_{t+1}^{nj}\dot{L}_{t+1}^{nj} w_t^{nj} L_t^{nj} = \gamma^{nj}(1-\xi^n)\sum_i \pi_{t+1}^{ij,nj} X_{t+1}^{ij},$$

where $\chi_{t+1} = \sum_i \sum_k \frac{\xi^i}{1-\xi^i} \dot{w}^{ik}_{t+1} \dot{L}^{ik}_{t+1} w^{ik}_t L^{ik}_t$ *and* $\dot{w}^n_t$ *is a vector whose* j*th element is* $\dot{w}^{nj}_t$*.*

Proof. The last five equations write the equilibrium conditions for the static multicountry interregional trade model in time differences. See CDP for a proof of this representation. Note that the real wage in period t can be written as

$$\frac{w^{nj}_t}{P^n_t} = \frac{\dot{w}^{nj}_t}{\dot{P}^n_t} \cdot \frac{\dot{w}^{nj}_{t-1}}{\dot{P}^n_{t-1}} \cdot \cdots \cdot \frac{\dot{w}^{nj}_1}{\dot{P}^n_1} \frac{w^{nj}_0}{P^n_0}.$$

Since the economy is initially starting from a stationary equilibrium at period $t = 0$, we have

$$\mathrm{d}\ln\left(\frac{w^{nj}_t}{P^n_t}\right) = \ln\left(\frac{\dot{w}^{nj}_t}{\dot{P}^n_t} \cdot \frac{\dot{w}^{nj}_{t-1}}{\dot{P}^n_{t-1}} \cdot \cdots \cdot \frac{\dot{w}^{nj}_1}{\dot{P}^n_1}\right). \tag{OA.16}$$

Note that with shocks known to agent at period $t = 1$, the sufficient statistics results of Proposition 1 can be simplified to

$$\mathrm{d}v_t = \sum_{k=0}^{\infty} \beta^k \mathcal{F}_k \,\mathrm{d}w_{t+k}$$

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

Plugging expression (OA.16) into these equations gives the first two equations. □

In the baseline economy, the path of exogenous state variables are given by $\{\Theta_t\}^{\infty}_{t=0}$ and $\bar{\Theta}$. In the counterfactual economy, we consider changes in exogenous state variables. We denote the new path by $\{\Theta'_t\}^{\infty}_{t=0}$. We assume that agents learn about these changes at period $t = 1$. This proposition corresponds to Proposition 3 of CDP, but allowing for worker heterogeneity. It shows how to solve for the counterfactual changes in endogenous variables in time differences and relative to a baseline economy without the need to estimate the level of the exogenous state variables. We denote by $\hat{y}_{t+1} \equiv \dot{y}'_{t+1}/\dot{y}_{t+1}$ the ratio of time differences between the counterfactual equilibrium and the baseline equilibrium.

Proposition OA.5 (Solving for Counterfactuals). *Suppose that the economy is initially starting from a stationary equilibrium at period* $t = 0$*. Up to the first-order approximation around a stationary equilibrium, given a baseline equilibrium,* $\{L_t, \pi_t, X_t\}^{\infty}_{t=0}$*, and a counterfactual sequence of changes in exogenous state variables,* $\{\hat{\Theta}_t\}^{\infty}_{t=1}$ *satisfying* $\lim_{t\to\infty} \hat{\Theta}_t = 1$*, known to agents in period* $t = 1$*, the solution to the counterfactual sequential equilibrium in time differences does not require information on the level of the exogenous state variables* $\{\Theta_t\}^{\infty}_{t=0}$ *or* $\bar{\Theta}$*, and solves the following system of equations:*

$$v'^{nj}_t = v^{nj}_t + \sum_k \beta^k \mathcal{F}_k \ln\left(\frac{\hat{w}^n_{t+k}}{\hat{P}^n_{t+k}} \cdot \frac{\hat{w}^n_{t+k-1}}{\hat{P}^n_{t+k-1}} \cdot \cdots \cdot \frac{\hat{w}^n_1}{\hat{P}^n_1}\right),$$

$$L'^{nj}_t = L^{nj}_t \cdot \exp\left(\left(\sum_{s=0}^{t-2} \sum_{k=0}^{\infty} \frac{\beta^{k+1}}{\rho} (\mathcal{F}^n_{s+k} - \mathcal{F}^n_{s+k+2}) \ln\left(\frac{\hat{w}^n_{t-s+k}}{\hat{P}^n_{t-s+k}} \cdot \frac{\hat{w}^n_{t-s+k-1}}{\hat{P}^n_{t-s+k-1}} \cdot \cdots \cdot \frac{\hat{w}^n_1}{\hat{P}^n_1}\right)\right)_j\right),$$

$$\hat{x}^{nj}_{t+1} = (\hat{L}^{nj}_{t+1})^{\gamma^{nj}\xi^n} (\hat{w}^{nj}_{t+1})^{\gamma^{nj}} \prod_k (\hat{P}^{nk}_{t+1})^{\gamma^{nj,nk}},$$

$$\hat{P}^{nj}_{t+1} = \left(\sum_i \pi'^{nj,ij}_t \dot{\pi}^{nj,ij}_{t+1} (\hat{x}^{ij}_{t+1} \hat{\kappa}^{nj,ij}_{t+1})^{-\theta^j} (\hat{A}^{ij}_{t+1})^{\theta^j \gamma^{ij}} \right)^{-1/\theta^j},$$

$$\pi'^{nj,ij}_{t+1} = \pi'^{nj,ij}_t \dot{\pi}'^{nj,ij}_{t+1} \left(\frac{\hat{x}^{ij}_{t+1} \hat{\kappa}^{nj,ij}_{t+1}}{\hat{P}^{nj}_{t+1}} \right)^{-\theta^j} (\hat{A}^{ij}_{t+1})^{\theta^j \gamma^{ij}},$$

$$X'^{nj}_{t+1} = \sum_k \gamma^{nk,nj} \sum_i \pi'^{ik,nk}_{t+1} X'^{ik}_{t+1} + \alpha^j \left(\sum_k \hat{w}^{nk}_{t+1} \hat{L}^{nk}_{t+1} w'^{nk}_t L'^{nk}_t \dot{w}^{nk}_{t+1} \dot{L}^{nk}_{t+1} + \iota^n \chi'_{t+1} \right),$$

$$\hat{w}^{nj}_{t+1} \hat{L}^{nj}_{t+1} = \frac{\gamma^{nj}(1-\xi^n)}{w'^{nk}_t L'^{nk}_t \dot{w}^{nk}_{t+1} \dot{L}^{nk}_{t+1}} \sum_i \pi'^{ij,nj}_{t+1} X'^{ij}_{t+1},$$

where $\chi'_{t+1} = \sum_i \sum_k \frac{\xi^i}{1-\xi^i} \hat{w}^{ik}_{t+1} \hat{L}^{ik}_{t+1} w'^{ik}_t L'^{ik}_t \dot{w}^{ik}_{t+1} \dot{L}^{ik}_{t+1}$ *and* $\hat{w}^n_t$ *is a vector whose* j*th element is* $\hat{w}^{nj}_t$.

OA.3.2.3 Calibration of the Model

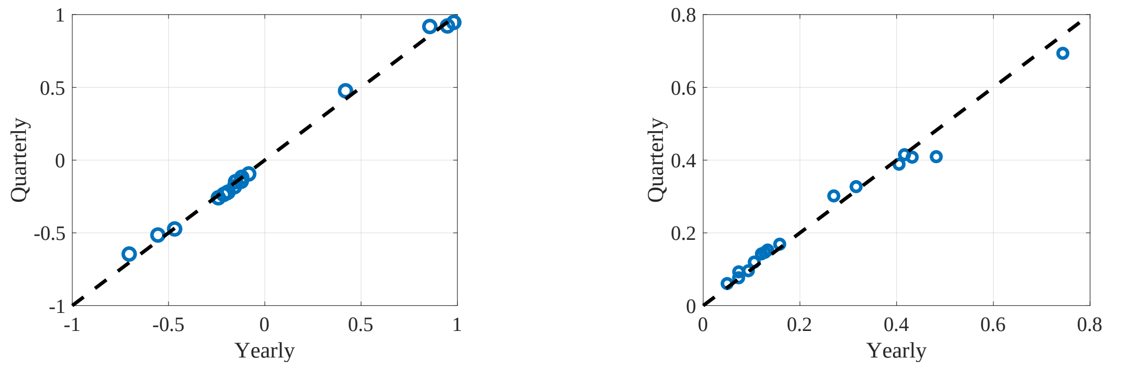
There are 87 regions, 50 US states and 37 other countries, and 4 sectors. The model has the following parameters: value added shares ($\{\gamma^{nj}\}_{n,j}$), the share of structures in value added ($\{\xi^n\}_n$), the input-output coefficients ($\{\gamma^{nk,nj}\}_{n,k,j}$), rentier shares ($\{\iota^n\}_n$), consumption Cobb-Douglas shares ($\{\alpha^j\}_j$), the discount factor (β), the sectoral trade elasticities ($\{\theta^j\}_j$), and the inverse sector-choice elasticity (ρ).[5] The year 2000 corresponds to the period $t = 0$ of the model. To apply dynamic hat algebra, we use data on bilateral trade flows π_t and value added $\{w^{nj}_t L^{nj}_t + r^{nj}_t H^{nj}\}_{n,j}$ from year 2000 to 2007. The data comes from the World Input-Output Database (WIOD), the 2002 Commodity Flow Survey (CFS), and regional employment data from the Bureau of Economic Analysis (BEA). See CDP for more details. Finally, we need to identify the magnitude of the China shock.

Parameters. Following CDP, value added shares ($\{\gamma^{nj}\}_{n,j}$), the share of structures in value added ($\{\xi^n\}_n$), and the input-output coefficients ($\{\gamma^{nk,nj}\}_{n,k,j}$) are constructed from the BEA and the WIOD data. Rentier shares ($\{\iota^n\}_n$), consumption Cobb-Douglas shares ($\{\alpha^j\}_j$) are calculated from the constructed trade and production data. We set the quarterly discount factor to $\beta = 0.99$. We use the sectoral trade elasticities from Dix-Carneiro et al. (2023), $\theta^j = 4$. Finally, we obtain the inverse migration elasticity at a quarterly frequency from the estimate in Section 4.5. In particular, the value of ρ at a quarterly frequency is calibrated such that both the yearly and quarterly analysis deliver the same elasticity of labor with respect to the wage changes. Up to a first-order approximation, the response of labor to a permanent change in wage $w_t = w$ known to households at $t = 1$ is given by

$$\mathrm{d} \ln \ell_t^{\text{ quarterly}} = \sum_{s=0}^{t-2} \sum_{k=0}^{\infty} \frac{(\beta^{\text{quarterly}})^{k+1}}{\rho^{\text{quarterly}}} (\mathcal{F}^{\text{quarterly}}_{s+k} - \mathcal{F}^{\text{quarterly}}_{s+k+2}) \, \mathrm{d}w,$$

$$\mathrm{d} \ln \ell_t^{\text{ yearly}} = \sum_{s=0}^{t-2} \sum_{k=0}^{\infty} \frac{(\beta^{\text{yearly}})^{k+1}}{\rho^{\text{yearly}}} (\mathcal{F}^{\text{yearly}}_{s+k} - \mathcal{F}^{\text{yearly}}_{s+k+2}) \, \mathrm{d}w.$$

[5] Without loss of generality, we can ignore $\{\eta^{nj}\}_{n,j}$ as they only appear in the constant term of the price index.



(a) Coefficients on Sectoral Employment Changes

(b) Coefficients on Sectoral Welfare Changes

Figure OA.5. Yearly to Quarterly

We calculate the value of $\rho^{\text{quarterly}}$ that minimizes the difference between $\frac{\mathrm{d}\ln \ell_2^{\text{yearly}}}{\mathrm{d}w}$ and $\frac{\mathrm{d}\ln \ell_2^{\text{quarterly}}}{\mathrm{d}w}$:

$$\left\| \sum_{k=0}^{\infty} \frac{(\beta^{\text{quarterly}})^{k+1}}{\rho^{\text{quarterly}}} (\mathcal{F}_k^{\text{quarterly}} - \mathcal{F}_{k+2}^{\text{quarterly}}) - \sum_{k=0}^{\infty} \frac{(\beta^{\text{yearly}})^{k+1}}{\rho^{\text{yearly}}} (\mathcal{F}_k^{\text{yearly}} - \mathcal{F}_{k+2}^{\text{yearly}}) \right\|_2 .$$

The resulting value is $\rho = 1.0011$. Figure OA.5a plots the elements of $\frac{\mathrm{d}\ln \ell_2^{\text{yearly}}}{\mathrm{d}w}$ against the correponding elements of $\frac{\mathrm{d}\ln \ell_2^{\text{quarterly}}}{\mathrm{d}w}$. In Figure OA.5b, we also compare the (normalized) sectoral value changes at the yearly and quarterly frequency:

$$(1-\beta^{\text{quarterly}})\,\mathrm{d}v_0 = (1-\beta^{\text{quarterly}}) \sum_{k=0}^{\infty} (\beta^{\text{quarterly}})^k \mathcal{F}_k^{\text{quarterly}}\,\mathrm{d}w$$

$$(1-\beta^{\text{yearly}})\,\mathrm{d}v_0 = (1-\beta^{\text{yearly}}) \sum_{k=0}^{\infty} (\beta^{\text{yearly}})^k \mathcal{F}_k^{\text{yearly}}\,\mathrm{d}w.$$

China Shock. Following CDP, we first compute the predicted increases in US imports from China between 2000 and 2007 using the increases in imports from China of other eight advanced economies during the same period as an instrument. Given this plausibly China-driven increase in imports, we calibrate the increase in China's manufacturing TFP, $\hat{A}_t^{\text{China,manufacturing}}$, from 2000 to 2007 such that the increase in imports from China driven by this increase in China's TFP exactly matches the predicted imports increase.

OA.4 Omitted Proofs

Proof of Proposition OA.1. We can write the absolute changes in the value as

$$|\,\mathrm{d}v_{s1}| = \beta^{\tau-1}(\mathcal{F}_{\tau-1})_{s,s}\Delta \geq \beta^{\tau-1}\big((\mathcal{F}_1)^{\tau-1}\big)_{s,s}\Delta = \Big|\,\mathrm{d}v_{s1}\big|_{\text{canonical}}\Big|. \qquad \square$$

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

$$\begin{aligned}
\mathrm{d}\ln \ell_t &= \sum_{s\geq 0, k\geq 0} \frac{\beta^{k+1}}{\rho}(\mathcal{F}_{s+k} - \mathcal{F}_{s+k+2})\,\mathbb{E}_{t-s-1}\,\mathrm{d}w_{t-s+k} \\
&= \sum_{s=0}^{t-2}\sum_{k\geq 0} \frac{\beta^{k+1}}{\rho}(\mathcal{F}_{s+k} - \mathcal{F}_{s+k+2})\,\mathrm{d}w_{t-s+k} \\
&= \tfrac{\beta}{\rho}(\mathcal{F}_{t-2} - \mathcal{F}_t)\,\mathrm{d}w_2 \ + \Big(\tfrac{\beta}{\rho}(\mathcal{F}_{t-3} - \mathcal{F}_{t-1}) + \tfrac{\beta^2}{\rho}(\mathcal{F}_{t-1} - \mathcal{F}_{t+1})\Big)\,\mathrm{d}w_3 \\
&\quad + \Big(\tfrac{\beta}{\rho}(\mathcal{F}_{t-4} - \mathcal{F}_{t-2}) + \tfrac{\beta^2}{\rho}(\mathcal{F}_{t-2} - \mathcal{F}_{t}) + \tfrac{\beta^3}{\rho}(\mathcal{F}_{t} - \mathcal{F}_{t+2})\Big)\,\mathrm{d}w_4 + \cdots \\
&\quad + \Big(\tfrac{\beta}{\rho}(\mathcal{F}_{0} - \mathcal{F}_{2}) + \cdots + \tfrac{\beta^{t-1}}{\rho}(\mathcal{F}_{2t-4} - \mathcal{F}_{2t-2})\Big)\,\mathrm{d}w_t \\
&\quad + \Big(\tfrac{\beta^2}{\rho}(\mathcal{F}_{1} - \mathcal{F}_{3}) + \cdots + \tfrac{\beta^{t}}{\rho}(\mathcal{F}_{2t-3} - \mathcal{F}_{2t-1})\Big)\,\mathrm{d}w_{t+1} + \cdots \equiv \sum_{\tau=2}^{\infty} \boldsymbol{A}_{\tau,t}\,\mathrm{d}w_\tau
\end{aligned}$$

where the impulse response functions are given by

$$\boldsymbol{A}_{\tau,t} = \begin{cases} \displaystyle\sum_{s=0}^{t-2} \frac{\beta^{s+\tau-t+1}}{\rho}(\mathcal{F}_{2s+\tau-t} - \mathcal{F}_{2s+\tau-t+2}) & \text{if } \tau \geq t \\ \displaystyle\sum_{s=0}^{\tau-2} \frac{\beta^{s+1}}{\rho}(\mathcal{F}_{2s+t-\tau} - \mathcal{F}_{2s+t-\tau+2}) & \text{if } t > \tau. \end{cases}$$

Thus, the s-th diagonal element of $\boldsymbol{A}_{\tau,t}$ is a weighted sum of $\{b_{2s+|t-\tau|}\}_{s=0,\ldots,t\wedge\tau-2}$ where more weights are given to those with small s. Thus, under Assumption 3, the s-th diagonal element of $\boldsymbol{A}_{\tau,t}$ is higher in the canonical model when $|t-\tau|$ and/or $t\wedge\tau$ are small. In particular, for given $t\wedge\tau$, we can find $B\in\mathbb{N}$ such that the s-th diagonal element of $\boldsymbol{A}_{\tau,t}$ is higher in the canonical model when $|t-\tau|\leq B$. We can show that $B\geq 1$:

$$\begin{aligned}
(\boldsymbol{A}_{t,t})_{ss} &= \sum_{s=0}^{t-2}\frac{\beta^{s+1}}{\rho}b_{2s} \\
&= \frac{\beta}{\rho} - \frac{\beta(1-\beta)}{\rho}(\mathcal{F}_2)_{ss} - \frac{\beta^2(1-\beta)}{\rho}(\mathcal{F}_4)_{ss} - \cdots - \frac{\beta^{t-2}(1-\beta)}{\rho}(\mathcal{F}_{2t-4})_{ss} - \frac{\beta^{m}}{\rho}(\mathcal{F}_{2t-2})_{ss},
\end{aligned}$$

$$\begin{aligned}
(\boldsymbol{A}_{t,t+1})_{ss} &= \sum_{s=0}^{t-2}\frac{\beta^{s+2}}{\rho}b_{2s+1} \\
&= \frac{\beta^2}{\rho}(\mathcal{F}_1)_{ss} - \frac{\beta^2(1-\beta)}{\rho}(\mathcal{F}_3)_{ss} - \frac{\beta^3(1-\beta)}{\rho}(\mathcal{F}_5)_{ss} - \cdots - \frac{\beta^{t-1}(1-\beta)}{\rho}(\mathcal{F}_{2t-3})_{ss} - \frac{\beta^{t}}{\rho}(\mathcal{F}_{2t-1})_{ss},
\end{aligned}$$

and

$$
\begin{aligned}
(\boldsymbol{A}_{t+1,t})_{ss} &= \sum_{s=0}^{t-2} \frac{\beta^{s+1}}{\rho} b_{2s+1} \\
&= \frac{\beta}{\rho}(\mathcal{F}_1)_{ss} - \frac{\beta(1-\beta)}{\rho}(\mathcal{F}_3)_{ss} - \frac{\beta^2(1-\beta)}{\rho}(\mathcal{F}_5)_{ss} - \cdots - \frac{\beta^{t-2}(1-\beta)}{\rho}(\mathcal{F}_{2t-3})_{ss} - \frac{\beta^m}{\rho}(\mathcal{F}_{2t-1})_{ss}.
\end{aligned}
$$

Thus, by Lemma 2, we can see that $(\boldsymbol{A}_{\tau,t})_{ss}$ is higher in the canonical model if $|t-\tau| \leq 1$. This proves that B maps $\mathbb{N}$ into itself. □

Proof of Lemma OA.1.

$$
\begin{aligned}
F_{\omega}^{k+1} &= \begin{pmatrix} 1-\alpha_\omega f^k & \alpha_\omega f^k \\ \beta_\omega f^k & 1-\beta_\omega f^k \end{pmatrix} \begin{pmatrix} \bar{\alpha}_\omega & \alpha_\omega \\ \beta_\omega & \bar{\beta}_\omega \end{pmatrix} \\
&= \begin{pmatrix} \bar{\alpha}_\omega - \alpha_\omega \bar{\alpha}_\omega f^k + \alpha_\omega \beta_\omega f^k & \alpha_\omega(1-\alpha_\omega f^k + \bar{\beta}_\omega f^k) \\ \beta_\omega(\bar{\alpha}_\omega f^k + 1 - \beta_\omega f^k) & \alpha_\omega \beta_\omega f^k + \bar{\beta}_\omega - \beta_\omega \bar{\beta}_\omega f^k \end{pmatrix} \\
&= \begin{pmatrix} 1-\alpha_\omega f^{k+1} & \alpha_\omega f^{k+1} \\ \beta_\omega f^{k+1} & 1-\beta_\omega f^{k+1} \end{pmatrix}.
\end{aligned}
$$

Thus, we have $f^{k+1}(x) = (x-1)\cdot f^k(x) + 1$. Define $g^k(x) = x \cdot f^k(2-x) - 1$, then we have

$$
\begin{aligned}
g^{k+1}(x) &= x \cdot f^{k+1}(2-x) - 1 \\
&= x \cdot \big((1-x)\cdot f^k(2-x) + 1\big) - 1 \\
&= (1-x)g^k(x)
\end{aligned}
$$

and $g^1(x) = -(1-x)$. Thus, $g^k(x) = -(1-x)^k$ and hence $f^k(x) = \frac{1+g^k(2-x)}{2-x} = \frac{1-(x-1)^k}{2-x}$. □

OA.5 Additional Figures

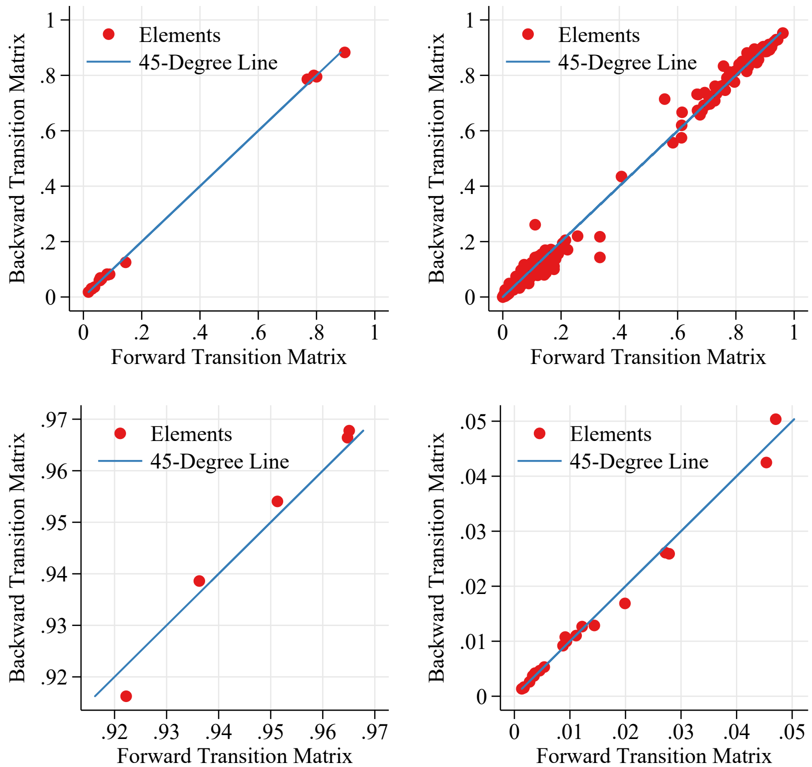


Figure OA.6. Backward and Forward Transition Matrices: NLSY (Top Row) and CPS (Bottom Row)

Notes: Assuming that the economy was in a steady state between years 1980 and 2000, the backward and forward transition matrices are computed by pooling all observations of the NLSY79 data over this period. In the left panel in the top row, we plot the elements of the aggregate forward transition matrix against those of the aggregate backward transition matrix. With four sectors, the backward and forward transition matrices are four-by-four matrices with sixteen elements. In the right panel, we consider four dimensions of observed heterogeneity—sex, race, education, and age, leading to sixteen groups—and compare the matrices for all sixteen groups. For CPS, we plot the aggregate backward and forward transition matrices in the bottom row. These matrices are close to the identity matrix, so we separately plot diagonal elements (left panel) and off-diagonal elements (right panel).

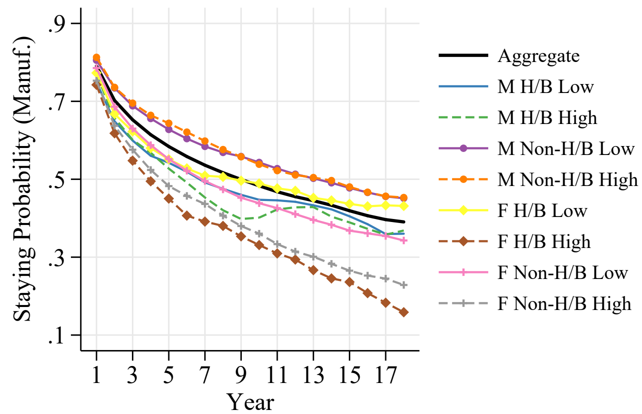
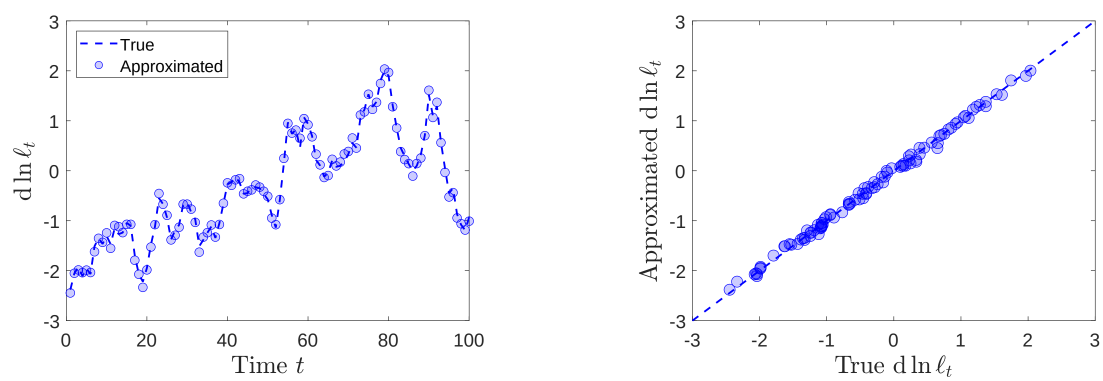


Figure OA.7. Actual Staying Probabilities for Different Worker Groups

Notes: For each unique combination of (male, female), (Hispanic/Black, non-Hispanic/Black), and (low-skilled, high-skilled), this figure plots the steady-state k-year manufacturing staying probabilities, $\Pr(s_{t+k} = \text{manufacturing} | s_t = \text{manufacturing})$. Data source: NLSY79.



(a) (b)

Figure OA.8. Fit of Recursive Representation

Notes: Plugging randomly generated values of $\{\mathrm{d}w_t\}$ (from standard normal distribution) and worker flow matrices computed from the NLSY data (extrapolated using the structural model) into equation (15), we can generate a sequence of changes in sectoral employment $\{\mathrm{d}\ln \ell_t\}$. Using the computed values of $\{\Gamma_k\}_{k=1,\ldots 6}$ and $\{\Lambda_k\}_{k=1,\ldots 5}$ and equation (17) instead, we can also compute an approximated sequence of changes in sectoral employment. These figures compare the actual sequence with the approximated sequence.

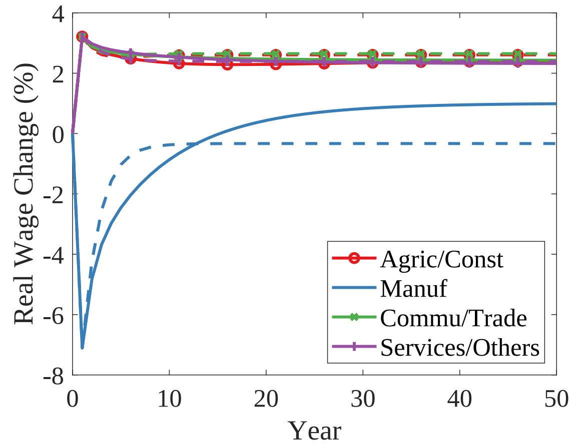


Figure OA.9. Changes in Sectoral Real Wages

Notes: This figure plots changes in sectoral real wages over time 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.

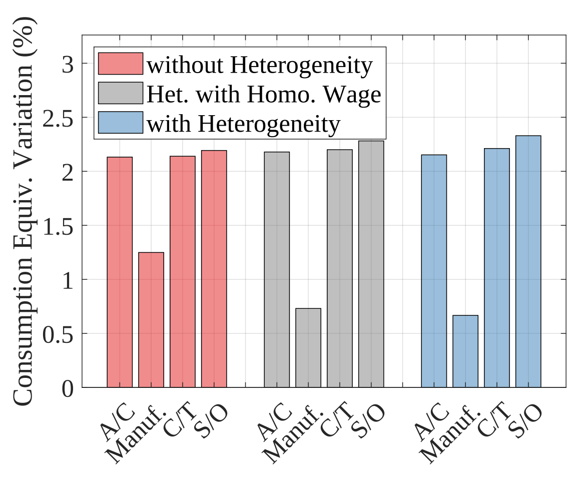
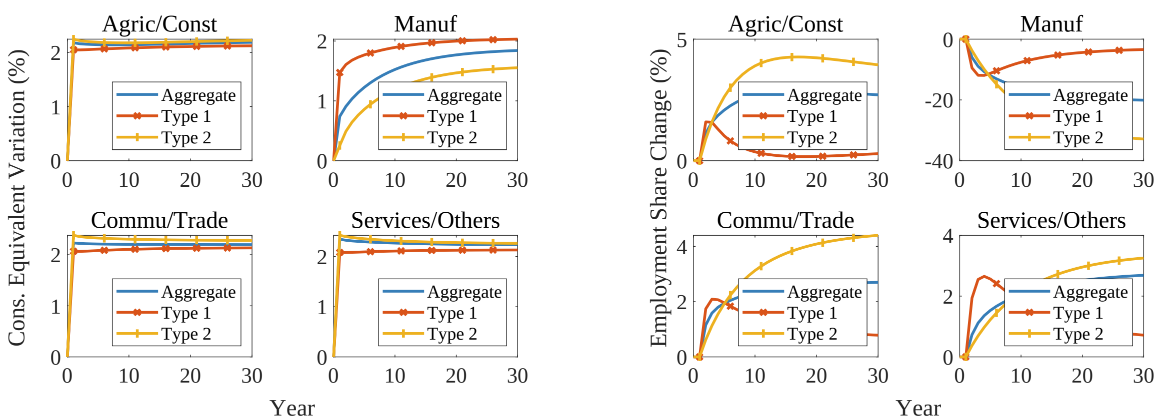


Figure OA.10. Changes in Sectoral Values: Exogenous and Endogenous Wage Changes

Notes: This figure plots changes in sectoral values in terms of consumption-equivalent variation for workers initially employed in different sectors. The last four bars correspond to the prediction from sufficient statistics in the data. The first four bars correspond to the prediction of the canonical model. The remaining bars in the middle correspond to the prediction made by combining the wage changes of the canonical model and the sufficient statistics.



(a) Changes in Sectoral Values

(b) Change in Sectoral Employment Share

Figure OA.11. Counterfactual Changes in Employment Share and Welfare: Type-specific Changes

Notes: This figure plots the transitional dynamics following an unexpected permanent drop in manufacturing prices for each type of worker. The orange line corresponds to type-1 workers (frequent movers), and the yellow line corresponds to type-2 workers (infrequent movers).

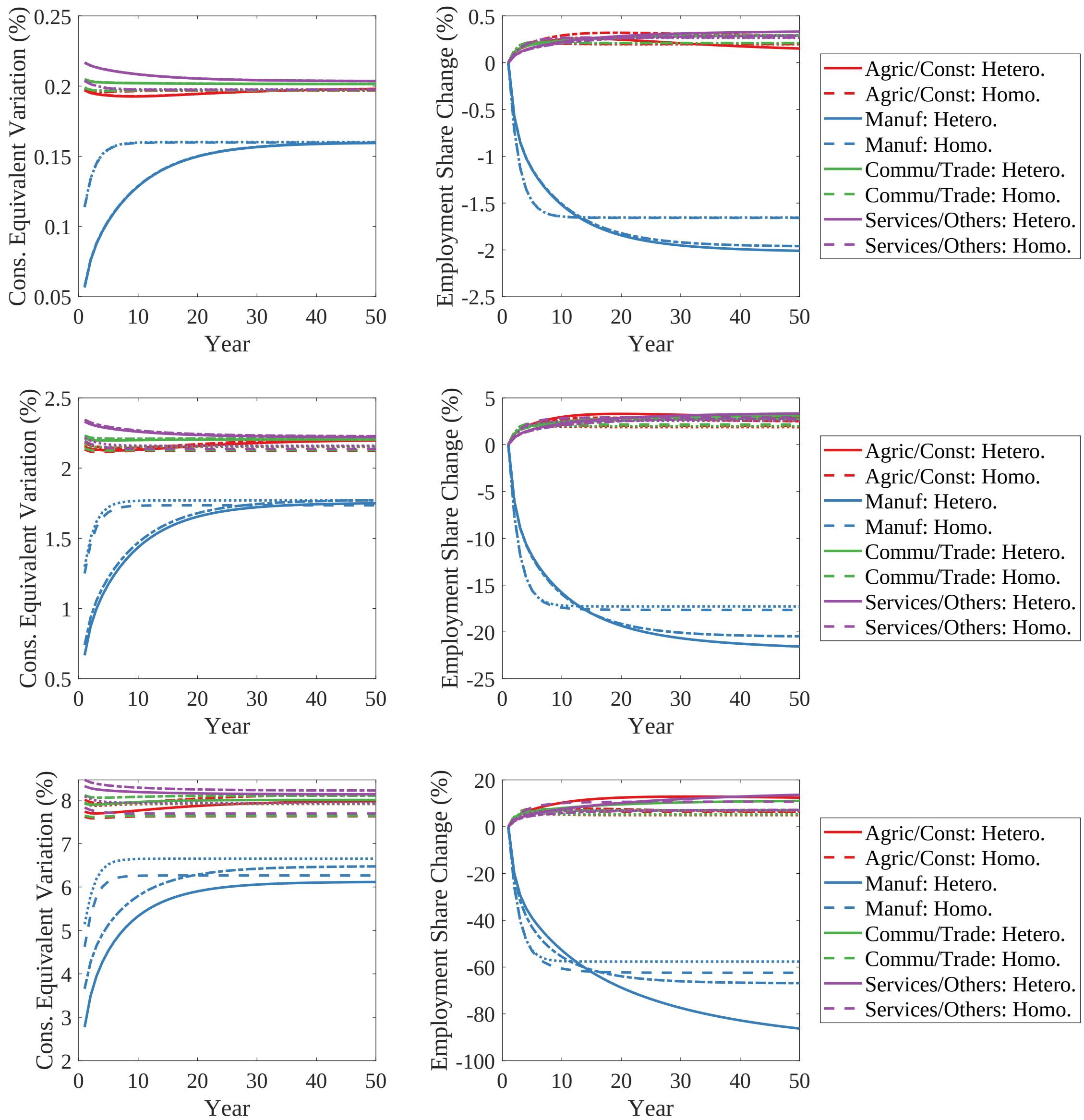


Figure OA.12. Quality of the First-Order Approximation

Notes: This figure compares the transitional dynamics of welfare (left column) and employment share (right column) obtained using sufficient statistics formula with those calculated from the exact solution of the estimated structural model. The top row corresponds to 1% drop in manufacturing prices, the middle row to 10%, and the bottom row to 30%.

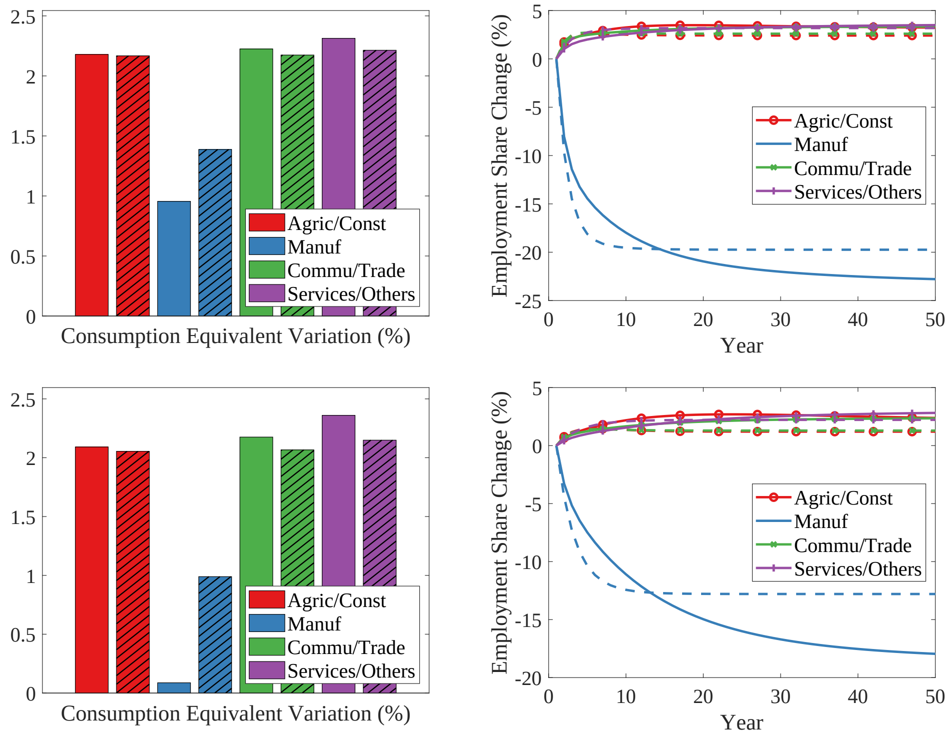


Figure OA.13. Counterfactual Exercises with Different Values of ρ

Notes: This figure reports the sectoral welfare change (left column) and the transitional dynamics of sectoral employment share (right column), computed with different values of ρ. The top row corresponds to $\rho = 0.5$, and the bottom row to $\rho = 2$.

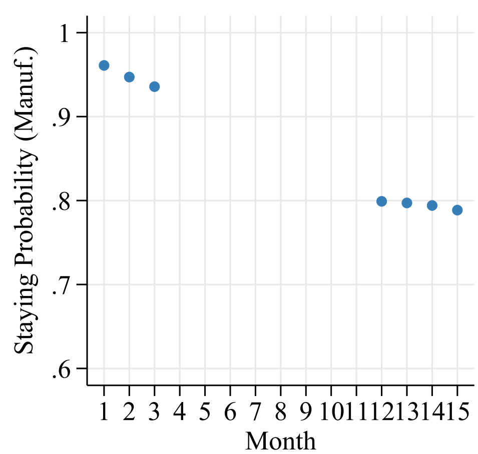
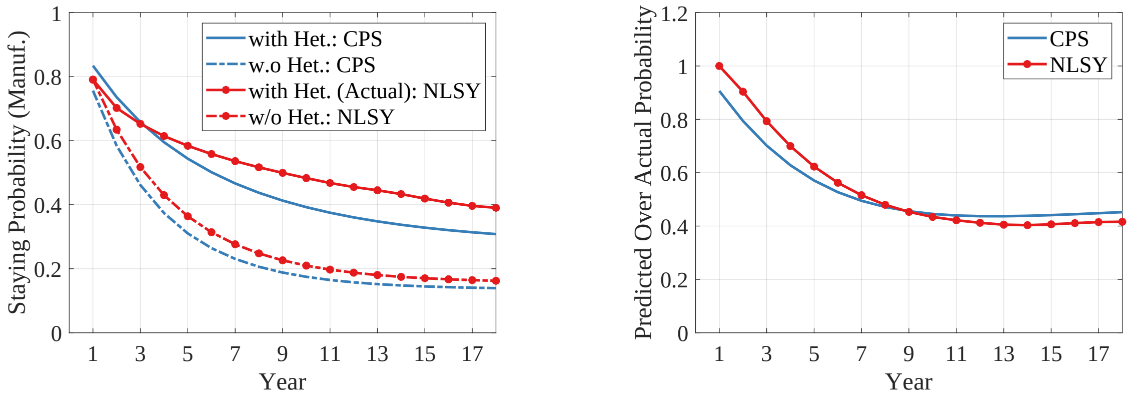


Figure OA.14. Aggregate Worker Flow Matrix Series: Monthly CPS

Notes: This figure clearly shows that the initial three circles are not in line with the remaining four circles. In particular, the last four circles should be shifted upwards. This is a well-known problem of the CPS dataset.



(a) Manufacturing: Actual and Predicted Probabilities (b) Manufacturing: Ratio between Actual and Predicted

Figure OA.15. Comparison: CPS and NLSY

Notes: In the left panel, the red lines depict the actual manufacturing staying probabilities from the NLSY79 data alongside those predicted by the canonical model. Using CPS data, we compute state-level worker flow matrices extrapolated by the two-type worker model and the canonical model. Aggregating these two series, the blue lines represent aggregate manufacturing staying probabilities implied by the models with and without worker heterogeneity. In the right panel, we calculate the ratio of these manufacturing staying probabilities, illustrating the degree to which the canonical model underestimates these probabilities.

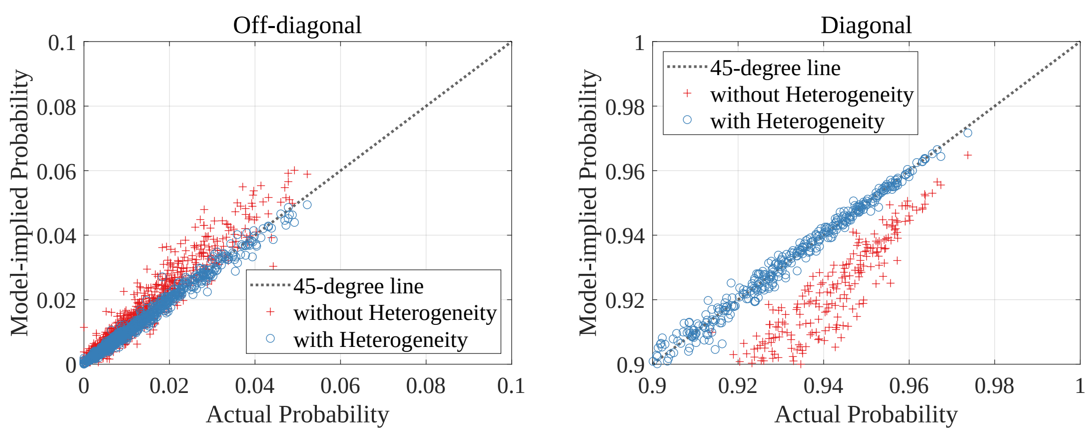
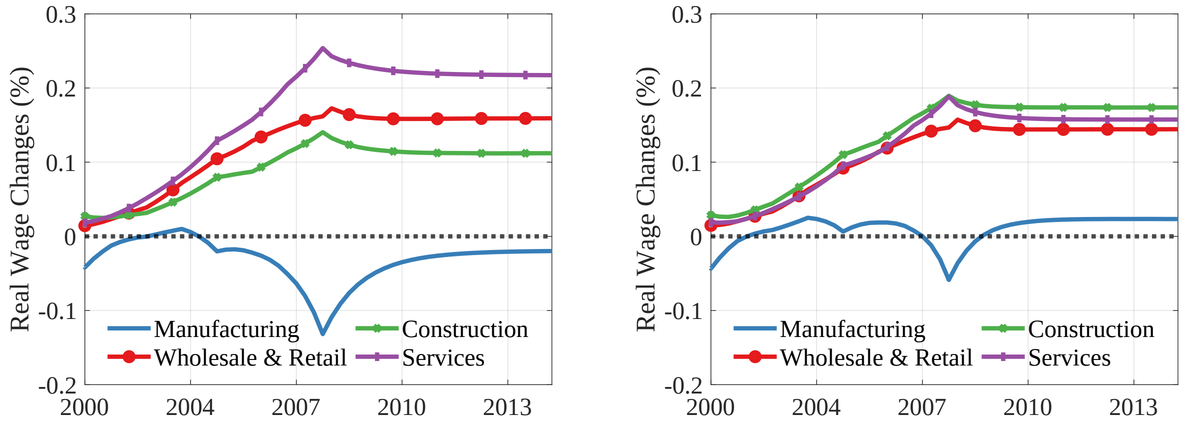


Figure OA.16. Fit of the Models with and without Heterogeneity: State-Level Worker Flow Matrices

Notes: In this figure, we plot the model-implied state-level worker flow matrix series for $k = 2, 3$ against that in the data. Blue circles represent the results of the heterogeneous-worker model and orange circles represent those of the canonical model, which exactly matches the 1-month worker flow matrix. Due to the short time horizon, worker flow matrices have diagonal elements close to one and off-diagonal elements close to zero. We separately plot them in the left and right panels.



(a) With Worker Heterogeneity

(b) Without Worker Heterogeneity

Figure OA.17. Changes in Sectoral Real Wages

Notes: This figure plots changes in sectoral real wages over time following the China shock.

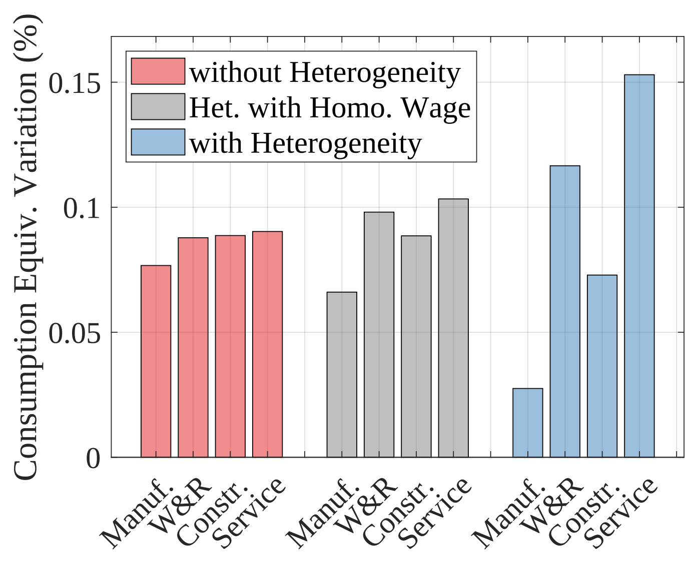


Figure OA.18. Changes in Sectoral Values: Exogenous and Endogenous Wage Changes

Notes: This figure plots changes in sectoral values in terms of consumption-equivalent variation for workers initially employed in different sectors. The last four bars correspond to the prediction from sufficient statistics in the data. The first four bars correspond to the prediction of the canonical model. The remaining bars in the middle correspond to the prediction made by combining the wage changes of the canonical model and the sufficient statistics.

References

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

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.

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

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.

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.

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.

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.

Henry, Neil W (1971) "The Retention Model: a Markov Chain with Variable Transition Probabilities," *Journal of the American Statistical Association*, 66 (334), 264–267.