NBER WORKING PAPER SERIES

# LEARNING AND EXPECTATIONS IN DYNAMIC SPATIAL ECONOMIES

Jingting Fan
Sungwan Hong
Fernando Parro

Working Paper 31504
http://www.nber.org/papers/w31504

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2023

We thank Lorenzo Caliendo, Jonathan Eaton, Yuhei Miyauchi, Nitya Pandalai-Nayar, Jonathan Vogel, and many seminar and conference participants for their useful conversations and comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Learning and Expectations in Dynamic Spatial Economies
Jingting Fan, Sungwan Hong, and Fernando Parro
NBER Working Paper No. 31504
July 2023
JEL No. F1,F11,F16,F18,Q54,R11,R13

**ABSTRACT**

The impact of shocks in dynamic environments depends on how forward-looking agents anticipate the path of future fundamentals that shape their decisions. We incorporate flexible beliefs about future fundamentals in a general class of dynamic spatial models, allowing beliefs to be evolving, uncertain, and heterogeneous across groups of agents. We show how to implement our methodology to study both ex-ante and ex-post shocks to fundamentals. We apply our method to two settings—an ex-ante study of the economic impacts of climate change, and an ex-post evaluation of the China productivity shock on the U.S. economy. In both cases, we study the impact of deviations from perfect foresight on different outcomes.

Jingting Fan
Department of Economics,
Pennsylvania State University
University Park, PA 16801
jxf524@psu.edu

Sungwan Hong
Pennsylvania State University
sph5642@psu.edu

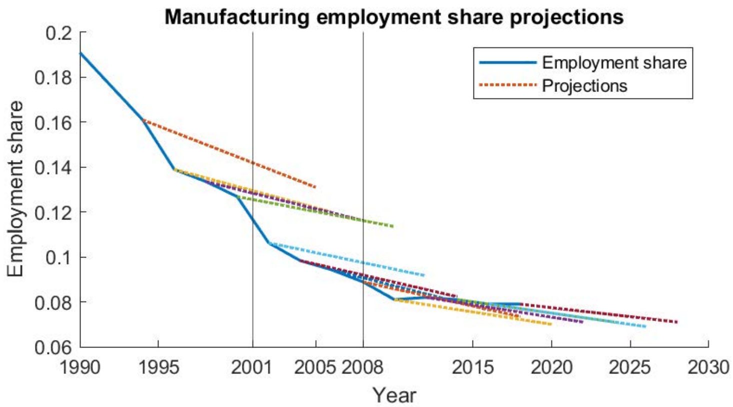
Fernando Parro
Pennsylvania State University
Department of Economics
University Park, PA 16801
and NBER
fxp5102@psu.edu

# 1 Introduction

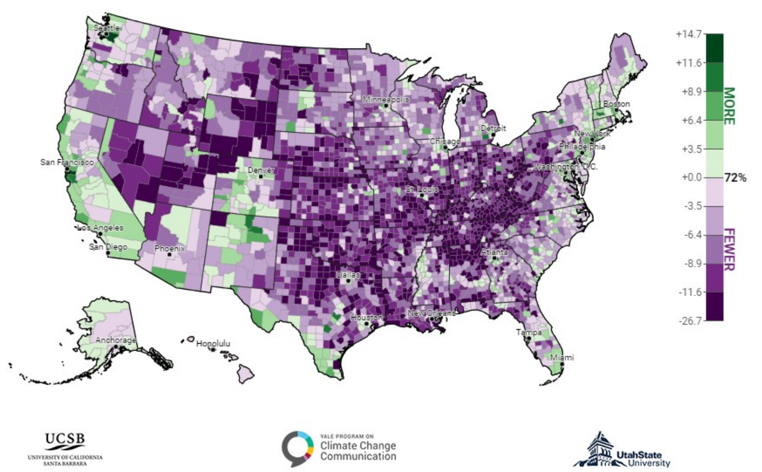
A central question in recent research is how changes in economic fundamentals shape the distribution of economic activity and impact different outcomes over time. The dynamic spatial effects of such changes are in part shaped by how forward-looking agents anticipate the trajectory of future fundamentals. For instance, if agents have perfect foresight, their decisions will be based on the actual path of fundamentals; if fundamentals are stochastic or if agents lack perfect foresight, then it is their beliefs about future fundamentals that shape their decisions; furthermore, if the beliefs differ across groups of agents, then agents take into account not only their own beliefs but also those of other groups since the decisions of different groups interact to determine the equilibrium outcomes. In this paper, we deviate from recent spatial dynamic models that assume perfect foresight. Our main contribution is to develop a methodology to study the spatial effects of shocks to fundamentals in a general stochastic dynamic spatial framework, allowing the beliefs about future fundamentals to be evolving, uncertain, or heterogeneous across groups of agents. We then apply our methodology to two settings—an ex-ante study of the economic impacts of climate change, and an ex-post evaluation of the China productivity shock on the U.S. economy.

To motivate our study, Figure 1a illustrates the relevance of departing from perfect foresight in the context of the China productivity shock. As the figure shows, manufacturing employment share in the United States was on a steady decline for a sustained period before it flattened out around 2008. Not everyone correctly anticipated the speed and persistence of the decline, however. The dotted lines plot the 10-year projection made by the Bureau of Labor Statistics, which started too optimistic but was revised downward over time, eventually converging to the actual outcomes by 2006-2008. To the extent that agents in the economy also did not foresee the declining trajectory—whether because they did not anticipate productivity growth in China or other changes in the economy—taking into account these beliefs could be important. Departing from perfect foresight is also relevant for understanding the impacts of climate change. Despite growing research on the subject, it remains highly uncertain how much the temperature around the world will rise over the next decades. In addition, even the very notion of climate change is not universally accepted—according to a recent survey, almost a third of adults in the United States do not believe climate change is happening. Figure 1b shows that these climate skeptics are unevenly distributed geographically, making heterogeneous beliefs important for understanding how climate change affects spatial outcomes.

To analyze the impacts of changes in fundamentals in stochastic environments like the ones described above, we set up a dynamic spatial model with forward-looking migration decisions and a general equilibrium trade structure across locations. Without loss of generality, we assume fundamental productivities across locations follow generic stochastic processes, which can be non-stationary and history-dependent, while other fundamentals are deterministic and can be time-varying. Agents make decisions based on their shared beliefs about these processes. The stochastic equilibrium of this model is characterized by an uncountable number of equilibrium conditions. The intuition is that forward-looking agents make mobility decisions based on the continuation




(a) U.S. Manufacturing Employment




(b) Share of Adults Who Think Global Warming is Happening

Figure 1: Motivation for Departing from Perfect Foresight Environments
Note: Panel (a) presents the manufacturing employment share in the United States during 1990-2018 (blue solid line) and 10-year manufacturing employment share projections made by the Bureau of Labor Statistics (dotted lines). Panel (b) presents the geographical distribution of difference from the national average (72%) in the share of adults who think global warming is happening in 2021. Source: Yale Climate Opinion Map (2021).

value of all locations, which in turn depends on the future utility flows in all possible future productivity realization. This process leads to an uncountable number of expected value functions one must solve for the actual migration decisions at each moment in time.

To make progress on solving the stochastic spatial equilibrium, we devise a local solution method that approximates the stochastic equilibrium around a deterministic transition path with perfect foresight. This approximation uses as inputs the deterministic transition path and the expectation of agents about future fundamentals. For any realization of the path of fundamentals, it gives as output the evolution of the spatial economy in deviations from the deterministic transition path—and this output has first-order accuracy. Compared to local approximations around steady states commonly used in the studies of business cycles, our approach has three advantages: it works in situations where the economy is non-stationary—for example, fundamentals (or agents' beliefs about these fundamentals) could have time trends, or the economy has a finite horizon, so a steady state does not exist; in settings where steady states exist, it is more accurate, especially if the economy is in the process of transitioning to a steady state with drastically different economic landscapes from the starting point due to either the large shock under investigation (e.g., climate change) or other forces. Last but not least, if the deterministic path is generated with time-varying fundamentals that are not the main focus of the counterfactual analysis—such as trade costs or migration costs in a study on climate change—then these fundamentals can be partialled out in our approximation. This will prove convenient in ex-post evaluations as we discuss below.

Building on this approach, we confront key aspects in quantifying the dynamic impact of shocks in stochastic spatial environments. In ex-post counterfactuals aimed to study the impact of a *past* event, researchers observe the outcomes under the actual realization of the fundamentals shaped by agents with evolving beliefs, but not the perfect foresight transition path corresponding to that realization. We propose an algorithm to recover the belief and perfect foresight paths from the observed allocations, allowing the underlying economy to have flexible beliefs as well as time-varying fundamentals such as trade and migration costs. Once the perfect foresight path has been recovered, we can approximate around it to conduct counterfactuals. Recognizing that the beliefs shaping the observed outcomes can be imperfect and carrying out the first step (as opposed to assuming the outcomes are generated by agents with perfect foresight) is important. For instance, in the context of the China shock, the lack of short-term migration response to import competition can be rationalized by either high migration costs or too pessimistic beliefs about China's productivity, which in turn can impact the subsequent counterfactual analysis.

We then extend our methodology to incorporate uncertainty in agents' beliefs. In particular, we extend our solution method to account for the variances-covariance of shocks and endogenous variables and show that it has the second-order precision that is needed to speak to the effect of uncertainty.

In addition, in environments with stochastic fundamentals, beliefs about the future might differ among groups of agents, as illustrated for the case of climate change in Figure 1b. In such settings, agents' decisions take into account not only their own beliefs but also the decisions of agents with different beliefs. We also extend our framework and methodology to incorporate heterogeneous beliefs. We show that our local solution method involves a curse of dimensionality problem akin to one in the literature of higher-order beliefs (Townsend, 1983). However, we propose alternative structures of heterogeneous beliefs that regain tractability, so an extension of the solution method described above applies.

We apply our methodology to two relevant settings. In the first application, we study the impact of rising temperatures over 2014-2100 on welfare and the spatial allocation of economic activity in the United States. We focus on the importance of two salient features of beliefs in this environment—that people are generally uncertain about future temperatures, and that a substantial share of the population does not believe in climate change (climate skeptics). As a baseline, we start by solving the sequential equilibrium when climate change is perfectly anticipated. Compared to a no climate change scenario, perfectly anticipated climate change leads to an average welfare decrease of 0.9% in the United States. Current residents in colder states, such as those in the Northeast or Midwest, benefit; those in the South lose. Such heterogeneous impacts also imply that the population gradually reallocates from the South to the North.

Incorporating uncertainty about future temperature rise increases welfare losses to 2% and leads to moderately faster reallocation across space. The intuition is that the marginal negative impact of temperature on productivity is higher where the temperature is higher. Because the South has a higher temperature today, an increase in the uncertainty in future temperature leads to a larger increase in the uncertainty in future productivity there, incentivizing people to move away faster. Separately from the first extension, we introduce climate skeptics into the model, with their initial distribution across states given by Figure 1b. We find that the presence of climate skeptics slows down spatial reallocation. It increases the welfare of other people (believers of climate change) in the North while decreasing the welfare of other people in the South.

Our second application is an ex-post evaluation of the China productivity catch-up over 2001-2008, in which agents have evolving and potentially biased beliefs. Using a model inversion, we recover the manufacturing productivity of China and the United States and estimate a process characterizing the catch-up of China's productivity. We assume that agents make decisions based on their beliefs about the catch-up process and discipline these beliefs using the 10-year manufacturing employment projections formed over 2001-2008 shown in Figure 1a.[1] We find that agents started out over-pessimistic about China's productivity, but these beliefs were revised and became aligned with reality by 2006. Through a counterfactual experiment, we find that the rapid productivity catch-up of China brings 1.1% welfare gains and leads to about a 0.3 million decrease in manufacturing employment between 2001 and 2008. If we had assumed data are generated under perfect foresight, we would have found a larger decrease in manufacturing employment by 2008 and slightly smaller welfare gains.

As mentioned above, our paper builds on recent spatial dynamic frameworks with perfect foresight environments, which have been used to fruitfully study questions such as the spatial effects of trade shock (e.g., Caliendo, Dvorkin and Parro, 2019; Rodríguez-Clare, Ulate and Vasquez, 2020; Dix-Carneiro, Pessoa, Reyes-Heroles and Traiberman, 2021), transitional dynamic with migration and capital accumulation (Kleinman, Liu and Redding, 2023), the impact of climate change across space (Desmet, Kopp, Kulp, Nagy, Oppenheimer, Rossi-Hansberg and Strauss, 2021, Cruz and Rossi-Hansberg, 2023; Balboni, 2019; Rudik, Lyn, Tan and Ortiz-Bobea, 2022), among other questions. Our crucial departure is to develop a spatial framework with stochastic fundamentals, and a methodology to study spatial effects of shocks to fundamentals in environments where the agent beliefs can be evolving, uncertain, or heterogeneous across groups of agents. Within this

[1] Our choice to focus on agents' beliefs about China's productivity is motivated by the recent studies on the role of imports from China in U.S. manufacturing employment (Autor et al., 2013; Pierce and Schott, 2016; Handley and Limão, 2017; Caliendo et al., 2019). Our methodology goes through if we assume instead that agents hold evolving beliefs about other fundamentals, such as manufacturing productivity in the United States.

strand of the literature, our local approximation approach is related to Kleinman, Liu and Redding (2023), who conduct first-order local approximation around steady states to extract analytical insights about the transition path in dynamic spatial models and to Bilal (2022), who develops a perturbation method for continuous-time dynamic spatial models. Relative to these works, we conduct local solutions around a transitional path with perfect foresight, which is more versatile in the way we described previously. In addition, our stochastic spatial framework, methodology, and applications accommodate various empirically relevant departures from perfect foresight, such as evolving, uncertain, and heterogeneous beliefs. In terms of the ability to incorporate aggregate uncertainty, our paper is related to Pang and Pin (2022), who study flood risk in a dynamic spatial model with aggregate uncertainty using a deep-learning approach. Relative to this paper, our approach allows for flexible aggregate shocks—for example, each location can have its own stochastic process, which can feature long history dependence—and we are able to prove its second-order accuracy.

By allowing agents to have flexible beliefs and by providing a method for inferring these beliefs, this paper is related to recent works by Dickstein and Morales (2018), Bombardini, Li and Trebbi (2023) and Fujiwara, Morales and Porcher (2020), who recover agents' beliefs based on their decisions. It differs in that our focus is on the general equilibrium outcomes. Our model of evolving beliefs is related to a growing macro and international literature emphasizing the role of learning, e.g., Cogley and Sargent (2005); Kozlowski, Veldkamp and Venkateswaran (2020); Bui, Huo, Levchenko and Pandalai-Nayar (2022). Aside from the focus on dynamic spatial models, this paper differs from most existing works in that it incorporates flexible beliefs updating processes about high-dimensional fundamentals in a non-stationary environment. We also provide a methodology to conduct ex-post counterfactuals based on the observed allocations.

Finally, our solution method is related to the second-order perturbation method in solving DSGE models pioneered by Schmitt-Grohé and Uribe (2004); Kim, Kim, Schaumburg and Sims (2008). The main difference is that while existing applications of such methods have a small number of aggregate states, our model accommodates a large number of states and sources of aggregate shocks—and flexible beliefs about these shocks, which makes it suitable for a wide range of questions that dynamic spatial models are designed for. Such tractability is achieved by exploiting the block-recursive property typical in a wide class of dynamic spatial models and the analytical tractability within each of these blocks.

The rest of the paper is structured as follows. In Section 2, we develop a dynamic stochastic spatial framework with flexible belief structures. In subsections 2.1 through 2.3 we describe the model. In subsection 2.4, we construct a local approximation method to undertake counterfactual analysis with evolving beliefs, and in subsection 2.5 we devise an algorithm for ex-post analysis. In subsections 2.6 and 2.7, we extend the framework and methodology to account for aggregate uncertainty and heterogeneous beliefs about the stochastic fundamentals. We then turn to our applications. In Section 3, we apply our framework and methodology to study the uncertainty and heterogeneous beliefs about climate change, and in Section 4 we study the effects of evolving beliefs about the China shock. Finally, Section 5 concludes. All proofs and detailed derivations are relegated to the appendix.

## 2 A Dynamic Stochastic Spatial Model

In this section, we develop a dynamic spatial model with stochastic fundamentals.

### 2.1 Economic Environment

We start by describing the economic environment. The world consists of $N$ locations, denoted by $n$ or $i$. Time $t$ is discrete and goes from 1 to $T$, which can be either a finite number or infinity. In each location, a continuum of individuals work and consume, and make forward-looking decisions on where to live next period subject to mobility frictions and idiosyncratic amenity preferences about locations. As described below, these dynamic migration decisions determine the supply of labor across locations at each moment in time. Such forward-looking labor supply decisions interact with a gravity trade structure to determine equilibrium wages and prices. To declutter the notations, we focus on a single-sector model in this section; all results generalize to a multi-sector extension with input-output linkages, which we will use in quantitative applications.

Each location $n$ is characterized by an initial endowment of labor $l_{n1}$ and by a set of potentially time-varying economic fundamentals $\{m_{nit}, \kappa_{nit}, z_{nt}\}$, where $m_{nit}$ and $\kappa_{nit}$ are bilateral mobility and trade frictions, respectively, and $z_{nt}$ are the location-specific fundamental productivity. Without loss of generality, we assume that productivity $z_{nt}$ evolves stochastically, and that trade and migration costs are characterized by a deterministic path perfectly anticipated by the agents. Our methodology extends to when $m_{nit}$ or $\kappa_{nit}$ are stochastic.

### 2.2 Stochastic Fundamentals and Beliefs

We now describe the evolution of fundamental productivity $z_{nt}$. We denote the *vector* of productivity across locations in period $t$ by $z_t \equiv (z_{1t}, z_{2t}, ..., z_{Nt})$ and the history of the productivity vector up to period $t$ by $z^t \equiv (z_1, z_2, ..., z_t)$. We denote the set of possible outcomes for $z_t$ in one period by $\Theta$ and the set of possible histories for $z^t$ by $\Theta^t$. For example, when $z_{nt}$ can take any positive values, $\Theta^t$ is $R_+^{N\times t}$.

Fundamental productivity $\boldsymbol{z}_t$ evolves according to a conditional probability density function (pdf) $g(z_{t+1}|z^t)$, where $g$ specifies the functional form as well as the parameters of the stochastic process governing $z_t$. This pdf characterizes the likelihood of all possible histories. For example, the likelihood of history $z^{t'} \in \Theta^{t'}$ for $t' > t$ is defined by $g(z^{t'}|z^t) \equiv \prod_{i=0}^{t'-t-1} g(z_{t+i+1}|z^{t+i})$.

Agents do not necessarily know $g(z_{t+1}|z^t)$. We denote by $f(z_{t+1}|z^t)$ the conditional pdf in the eyes of period-t agents after a history $z^t$, which we assume for now are common to all agents (later on we will extend the model to accommodate heterogeneous beliefs). It is under the guidance of $f(z_{t+1}|z^t)$ that agents make forward-looking decisions to maximize their utility.

Through the choice of $f$ and $g$, our model accommodates many setups. First, under the special case of $f = g$, agents have rational expectations. Second, both $g$ and $f$ can be non-stationary and time-varying, i.e., the process governing fundamentals today could be different from the process governing the fundamentals in the future. Without loss of generality, this time dependence is subsumed in the dependence of $f$ and $g$ on $z^t$. In the special case where $f$ and $g$ depend on $z_t$ instead of $z^t$, they become Markovian processes. Third, when $f$ differs from $g$, it can be either because agents are wrong about the functional form of the stochastic processes, or because they gradually learn about the parameters governing the processes.[2] Fourth, by specifying how $f$ varies with new information, our model captures different forms of learning. For example, agents could engage in Bayesian or myopic learning; regardless of the case, they could be 'naive' learners as defined by the anticipated utility framework developed by Kreps (1998), who think that their understanding of the world is correct and will not change in the future—only to find out that they were wrong; or they could be 'sophisticated' in the sense that they know in the future as new data arrive, they will revise their current beliefs.[3]

### 2.3 Dynamic Forward-Looking Migration Decisions with Stochastic Fundamentals

The continuum of agents residing in the economy consume in their current location and make forward-looking decisions on where to live subject to mobility frictions $m_{nit}$ and idiosyncratic taste shocks $\epsilon_{it}$. We assume $\epsilon_{it}$ are i.i.d realizations of a Gumbel (Type-I Extreme Value Distribution) with dispersion parameter $\nu$ (e.g. Artuç et al., 2010; Caliendo et al., 2019), and are orthogonal to the stochastic fundamental productivities.

Formally, after the realization of history $z^t$, the value of a location $n$ for an agent with a realization of idiosyncratic taste shocks $\epsilon_t \equiv (\epsilon_{it}...\epsilon_{Nt})$ is given by

$$V_{nt}(z^t, \epsilon_t) = U(c_{nt}(z^t)) + \max_{\{i\}_{i=1}^N} \{\beta \mathbb{E}[V_{it+1}(z^{t+1}, \epsilon_{t+1})|z^t] - m_{nit} + \nu\epsilon_{it}\}, \tag{1}$$

where $U(c_{nt}(z^t))$ is the flow utility of the agent at location $n$ given the history $z^t$, with $c_{nt}(z^t)$ being the real income defined as the ratio between wage $w_{nt}(z^t)$ and the price of consumption goods $P_{nt}(z^t)$. The expectation term $\mathbb{E}[V_{it+1}(z^{t+1}, \epsilon_{it+1})|z^t]$ is the conditional expectation over both the idiosyncratic taste shocks $\epsilon_{t+1}$ and the possible realization of fundamental productivities at time $t+1$, defined by

$$\mathbb{E}[V_{it+1}(z^{t+1}, \epsilon_{it})|z^t] = \int_\Theta \left[ \int_{\epsilon_{t+1}} V_{it+1}(z^{t+1}, \epsilon_{t+1}) dH(\epsilon_{t+1}) \right] f(z_{t+1}|z^t) dz_{t+1.}$$

To reduce notation, we denote by $v_{it}(z^t)$ the expected value in period $t$ in location $i$, where the expectation is taking over the idiosyncratic shocks, namely $v_{it}(z^t) \equiv \int_{\epsilon_t} V_{it}(z^t, \epsilon_t) dH(\epsilon_t)$. Under

[2] To be more specific, we can separate the functional form from the parameters of the stochastic process. More concretely, if we denote by $\phi$ the parameters, then $g(z_{t+1}|z^t, \phi)$, and the agents perceive process is given by $f(z_{t+1}|z^t, \hat{\phi})$. Agents' beliefs might differ from the true process due to either $f \neq g$ or $\hat{\phi} \neq \phi$.

[3] Following the notation in Footnote 2, with naive agents we have that $f_t(z^{t'}|z^t, \hat{\phi}) = \prod_{i=0}^{t'-t-1} f(z_{t+i+1}|z^{t+i}, \hat{\phi}(z^0))$ and with sophisticated agents, the pdf of agent beliefs is $f_t(z^{t'}|z^t, \hat{\phi}) = \prod_{i=0}^{t'-t-1} f(z_{t+i+1}|z^{t+i}, \hat{\phi}(z^t))$.

this notation, the above equation can be written as

$$\mathbb{E}[V_{it+1}(z^{t+1}, \epsilon_{it})|z^t] = \int_{\Theta} v_{it+1}(z^{t+1}) f(z_{t+1}|z^t) dz_{t+1} \equiv \mathbb{E}[v_{it+1}(z^{t+1})|z^t].$$

The expectation of the value of location $n$ ($V_{nt}(z^t, \epsilon_t)$ defined in (1)) over different idiosyncratic draws is then:

$$v_{nt}(z^t) = U(c_{nt}(z^t)) + \nu \log \left( \sum_{i=1}^{N} \exp \left( \beta \mathbb{E}[v_{it+1}(z^{t+1})|z^t] - m_{nit} \right)^{1/\nu} \right), \tag{2}$$

where we use the properties of the Gumbel distribution to derive the continuation value. These properties also allow us to derive $\mu_{nit}(z^t)$, the gross migration flows from $n$ to $i$ in period $t$ after a history $z^t$:

$$\mu_{nit}(z^t) = \frac{\exp\left(\beta \mathbb{E}[v_{it+1}(z^{t+1})|z^t] - m_{nit}\right)^{1/\nu}}{\sum_{h=1}^{N} \exp\left(\beta \mathbb{E}[v_{ht+1}(z^{t+1})|z^t] - m_{nht}\right)^{1/\nu}}. \tag{3}$$

The migration decisions at time $t$ after a history $z^t$, together with the distribution of workers at the beginning of period $t$, which is determined in period $t-1$ and denoted by $l_t^n(z^{t-1})$, govern the distribution of individuals in $t+1$, namely,

$$l_{nt+1}(z^t) = \sum_{i=1}^{N} \mu_{int}(z^t) l_{it}(z^{t-1}). \tag{4}$$

Therefore, the equilibrium conditions (2), (3), and (4) determine the evolution of the distribution of labor supply across locations in this stochastic environment.

We model the demand side of the model with a general equilibrium gravity trade structure with CES demand over differentiated goods (e.g. Eaton and Kortum, 2002). The assumed structure implies that, in equilibrium, the share of goods purchased by location $n$ from location $i$, denoted by $\lambda_{nit}(z^t)$, is

$$\lambda_{nit}(z^t) = z_{it} \left( \frac{w_{it}(z^t)\kappa_{nit}}{P_{nt}(z^t)} \right)^{-\theta}, \tag{5}$$

where $w_{it}$ is the wage in location $i$ and $P_{nt}$ is the price index in location $n$ given by

$$P_{nt}(z^t) = \left[ \sum_{i=1}^{N} z_{it} \left( w_{it}(z^t)\kappa_{nit} \right)^{-\theta} \right]^{-1/\theta}.[4] \tag{6}$$

Finally, the labor market clearing condition is given by

$$w_{nt}(z^t) l_{nt}(z^t) = \sum_{i=1}^{N} \lambda_{int}(z^t) w_{it}(z^t) l_{it}(z^t). \tag{7}$$

Given the allocation of labor, equations (5), (6), and (7) solve the static equilibrium of the trade

[4] There is a constant in the price index in the Eaton and Korum formulation, which we omit without loss of generality.

model. We now define the stochastic sequential equilibrium of the economy.

**Definition 1.** ***A stochastic sequential equilibrium of the model*** *is a set of state-contingent prices* $\{w_{nt}(z^t), P_{nt}(z^t)\}_{n=1,t=1}^{N,T}$*, allocations of goods and labor* $\{\mu_{nit}(z^t), \lambda_{nit}(z^t), l_{nt}(z^{t-1})\}_{n=1,i=1,t=1}^{N,N,T}$*, and the value of locations* $\{v_{nt}(z^t)\}_{n=1,t=1}^{N,T}$ *that satisfies the equilibrium conditions determined by the location value function* (2)*, the gross flows equation* (3)*, the law of motion of labor* (4)*, the bilateral trade shares* (5)*, the local prices* (6)*, and the labor market clearing condition* (7)*, and are consistent with beliefs* $f(z_{t+1}|z^t)$*.*

Two remarks are in order. First, unlike in many other settings where the aggregate state of the economy at any point in time is summarized by the contemporary productivity $z_t$ and labor allocation $\{l_{it}\}_{i=1}^N$, in this definition, the state of the economy is indexed by the history of productivity until that point $z^t$. The time dependence in our model stems from the fact that $f$ could have history dependence, which enables the model to accommodate many belief-updating processes, and that $\{\kappa_{nit}, m_{nit}\}$ can vary arbitrarily over time. Only in the special case where $f$ is stationary and all fundamentals are constant can the aggregate state be summarized by $z_t$ and $\{l_{it}\}_{i=1}^N$.

Second, as all equations are defined on $z^t$, the equilibrium is characterized by an uncountable set of equations. The presence of forward-looking decisions implies that even solving for the decision rule for period one, one must solve the full set of equations for all period $t$ and all possible trajectories $z^t$. To see this aspect of the stochastic equilibrium more clearly, in equation (2), the value of each location given a history of fundamental productivities, $v_{nt}(z^t)$, depends on $\mathbb{E}[v_{it+1}(z^{t+1})|z^t]$, defined as

$$\begin{aligned}
&\mathbb{E}[v_{it+1}(z^{t+1})|z^t] \qquad (8)\\
&= \int_{\Theta} \left[ U\left(c_{it+1}(z^{t+1})\right) + \nu \log\left( \sum_{h=1}^{N} \exp\left(\beta \mathbb{E}[v_{ht+2}(z^{t+2})|z^{t+1}] - m_{iht+1}\right)^{1/\nu} \right) \right] f(z_{t+1}|z^t) dz_{t+1}.
\end{aligned}$$

The expected continuation value $\mathbb{E}[v_{it+1}(z^{t+1})|z^t]$ thus depends on the expected future utility flows and the future expected value in all $z^{t+1}$, which in turn is linked to outcomes in $z^{t+2}$ through the future expected values. This dependence continues until the end of the model, so solving the model exactly for the general setup is intractable. In what follows, we propose a local solution method of the stochastic spatial equilibrium around a deterministic transition path with perfect foresight. As we will show, our solution method extends tractably to the second order, so it can accommodate the role of uncertainty in agent behaviors; in addition, it can also be extended to incorporate heterogeneous beliefs among agents.

## 2.4 Local Approximation

We denote by $x_t(z^t)$ the equilibrium outcome of variable $x$ at time $t$ given a history up to $t$, denote by $\bar{x}_t$ the value of outcome $x$ in period $t$ in a deterministic path with perfect foresight, and denote by $\hat{x}_t(z^t)$ the deviation of $x_t(z^t)$ from $\bar{x}_t$. These deviations will be level differences for the value of locations and log differences for all other variables. For example, $\hat{z}_t$ denotes the deviation in the vector of fundamental productivity of period $t$ from the deterministic equilibrium $\log(z_t)$ −

$\log(\bar{z}_t)$. We define $v_{it+1}(z^t) \equiv \mathbb{E}[v_{it+1}(z^{t+1})|z^t]$ and define $v_{t+1}(z^t)$ as the vector obtained from stacking $v_{nt+1}(z^t)$ by location, namely $v_{t+1}(z^t) \equiv (v_{1t+1}(z^t), v_{2t+1}(z^t), ..., v_{Nt+1}(z^t))$. Then, under the hat notation, $\hat{v}_{t+1}(z^t) = v_{t+1}(z^t) - \bar{v}_{t+1}$.

Using these definitions, we derive $\hat{x}_t(z^t)$ as a function of $\bar{x}_t$ and $\hat{z}_t$. We start with $\hat{v}_{it+1}(z^t)$, the difference between equation (8) and its deterministic counterpart:

$$\begin{aligned}
\hat{v}_{it+1}(z^t) \equiv & \int_{\Theta} \left[ U\left(c_{it+1}(z^{t+1})\right) + \nu \log \left( \sum_{n=1}^{N} \exp\left(\beta v_{nt+2}(z^{t+1}) - m_{int+1}\right)^{1/\nu} \right) \right] f(z_{t+1}|z^t) dz_{t+1} \\
& - \left[ U\left(\bar{c}_{it+1}\right) + \nu \log \left( \sum_{n=1}^{N} \exp\left(\beta \bar{v}_{nt+2} - m_{int+1}\right)^{1/\nu} \right) \right] \\
\equiv & \int_{\Theta} \left[ F_i\left(c_{it+1}(z^{t+1}), v_{t+2}(z^{t+1})\right) - F_i\left(\bar{c}_{it+1}, \bar{v}_{t+2}\right) \right] f(z_{t+1}|z^t) dz_{t+1} \\
= & \int_{\Theta} \left[ \frac{\partial F_i}{\partial \bar{c}_{it+1}} \hat{c}_{it+1}(z^{t+1}) + \frac{\partial F_i}{\partial \bar{v}_{t+2}} \hat{v}_{t+2}(z^{t+1}) + o\left(\hat{v}_{t+2}(z^{t+1}), \hat{c}_{it+1}(z^{t+1})\right) \right] f(z_{t+1}|z^t) dz_{t+1},
\end{aligned}$$

where we have defined the terms inside the squared brackets in the first line as a function $F_i$, so $\hat{v}_{it+1}(z^t)$ can be seen as the difference between $F_i$ evaluated the stochastic path versus the deterministic path, integrated over possible realizations in $t+1$. The last line of the equation is a first-order Taylor approximation of $\left[F_i\left(c_{it+1}(z^{t+1}), v_{t+2}(z^{t+1})\right) - F_i\left(\bar{c}_{it+1}, \bar{v}_{t+2}\right)\right]$ at $(\bar{c}_{it+1}, \bar{v}_{t+2})$, with an approximation error $o(\hat{v}_{t+2}(z^{t+1}), \hat{c}_{it+1}(z^{t+1}))$ that is of second order to $\hat{v}_{t+2}(z^{t+1})$ and $\hat{c}_{it+1}(z^{t+1})$.[5] Integrating this equation over the possible path of fundamental productivities up to $t$ from the eyes of an agent making decisions at $t=1$, we obtain

$$\int_{\Theta^t} \hat{v}_{it+1}(z^t) f(z^t|z_1) dz^t \equiv \mathbb{E}_1 \hat{v}_{it+1} \approx \frac{\partial F_i}{\partial \bar{c}_{it+1}} \mathbb{E}_1 \hat{c}_{it+1} + \frac{\partial F_i}{\partial \bar{v}_{t+2}} \mathbb{E}_1 \hat{v}_{t+2}. \tag{9}$$

We show in the appendix that $\frac{\partial F_i}{\partial \bar{c}_{it+1}}$ and $\frac{\partial F_i}{\partial \bar{v}_{t+2}}$ are both closed-form functions of outcomes in the deterministic path, so we have obtained the expectation of the continuation value in period $t$ from the eyes of the agents in the first period, $\mathbb{E}_1 \hat{v}_{it+1}$, as a linear function of $\mathbb{E}_1 \hat{c}_{it+1}$ and $\mathbb{E}_1 \hat{v}_{t+2}$.

We can approximate the expected deviations in the gross mobility flows analogously. Specifically, express the mobility flows after a history of $z^t$ as

$$\log \mu_{nit}(z^t) = \log \frac{\exp\left(\beta v_{it+1}(z^t)] - m_{nit}\right)^{1/\nu}}{\sum_{h=1}^{N} \exp\left(\beta v_{ht+1}(z^t) - m_{nht}\right)^{1/\nu}} \equiv G_{ni}(v_{t+1}(z^t)).$$

Taking the first-order approximation of $G_{ni}(v_{t+1}(z^t)) - G_{ni}(\bar{v}_{t+1})$ around $\bar{v}_{t+1}$ and integrating the approximation across $z^t$ by their likelihood $f(z^t)$ gives us $\mathbb{E}_1 \hat{\mu}_{nit}$ as a linear function of $\mathbb{E}_1 \hat{v}_{t+1}$:

$$\mathbb{E}_1 \hat{\mu}_{nit} \approx \frac{\partial G_{ni}}{\partial \bar{v}_{t+1}} \mathbb{E}_1 \hat{v}_{t+1}.$$

[5]In $\frac{\partial F_i}{\partial \bar{x}}$, we use $\partial \bar{x}$ to denote level derivatives with respect to $x$ when x is the value of locations and log derivatives with respect to $x$ when $x$ is other variables. Under both cases, the derivatives are evaluated at the deterministic sequential equilibrium, hence the bar in $\partial \bar{x}$.

We apply similar procedures to the rest of the equilibrium conditions, which gives us a system of linear equations that takes as input $\mathbb{E}_1\hat{z}_t$ and gives as outputs the deviations in all endogenous outcomes from their deterministic counterparts.[6] The coefficients of all deviation terms in these equations (e.g., $\frac{\partial F_i}{\partial \bar{c}_{it+1}}$, $\frac{\partial F_i}{\partial \bar{v}_{t+2}}$, $\frac{\partial G_{ni}}{\partial \bar{v}_{t+1}}$) are closed-form expressions of the outcomes in the deterministic transition path with perfect foresight. Thus, we can solve this system of equations for the expected outcomes (expectation formed according to agents' belief in the first period) on the counterfactual path, and the solution takes the form of a linear function of $\mathbb{E}_1\hat{z}_t$. Along with the expected value for future outcomes, the solution to the system of the equations also gives us the actual outcomes in the first period. Importantly, as the approximation error in each of the linearized equations is second order in the expected deviations, and as the deviations themselves are expressed as linear functions of $\mathbb{E}_1\hat{z}_t$, the errors in these solutions are second order in $\mathbb{E}_1(\hat{z}_t)$.

Although our discussion thus far has focused on the decision (and expected outcomes) of the agents in the first period, it should be clear that such approximation can be applied to the decision of agents in any period $t$ after any history $z^t$. We summarize these in the following proposition.

**Proposition 1.** *Given the labor allocations in period $t$, agents' expectation at time $t$ about the deviations in endogenous variables at time $t' > t$, namely $\left\{\mathbb{E}_t\hat{v}_{it'}, \mathbb{E}_t\hat{\mu}_{nit'}, \mathbb{E}_t\hat{l}_{it'}, \mathbb{E}_t\hat{w}_{it'}, \mathbb{E}_t\hat{\lambda}_{nit'}, \mathbb{E}_t\hat{P}_{it'}\right\}_{i=1,n=1,t'=t}^{N,N,T}$, solve a system of linear equations (in the order of $N^2 \times (T-t)$), with the input being the outcomes on the deterministic (perfect foresight) path and the expected deviations of the fundamental productivities from that on the deterministic path $\mathbb{E}_t\hat{z}_{t'}$.*

*Proof.* See Appendix A.1. □

Proposition 1 shows that to obtain the decision of the agents in period $t$, instead of solving all equations in Definition 1, by approximating the nonlinear model around a deterministic transitional path starting at period $t$, we only need to solve a finite set of linear equations. In doing so, we also obtain the expected values for future outcomes according to agents' belief in period $t$, which we will call the *belief path* at $t$ for subsequent discussions. The input to the system of equations is a deterministic path with perfect foresight path from $t$ and the expected deviations of future fundamentals from the fundamentals underlying the deterministic path $\mathbb{E}(\hat{z}_{t'}|z^t), t' > t$. The perfect foresight path can be computed building on existing methods (e.g., Caliendo et al., 2019), although later on we address key aspects related to stochastic environments. The process that governs the shocks that the researcher is interested in studying and the agent's belief about the process can be guided and informed by data, as we illustrate in our second application.

Proposition 1 can be readily used to understand how agents' decisions in $t$ and expected outcomes for $t' > t$ change when their beliefs about future fundamentals deviate from perfect foresight. As a special case, when agents' beliefs are on average the same as the fundamentals underlying the deterministic path with perfect foresight, i.e., when $\mathbb{E}_t\hat{z}_{t'} = 0$, then up to the first order, agents' decision in $t$ and their expected value about future outcomes will be the same as the deterministic path—a familiar certainty equivalent result.

[6]The term $\mathbb{E}_1\hat{z}_t$ enters this system of equations through the static trade structure, see equations (5) and (6).

By applying Proposition 1 recursively, we can study how an economy evolves as agents change their beliefs over time along any given path of fundamentals $z^T$. Specifically, starting from the allocation in period $t$, one can solve for the perfect foresight path from $t$, and apply Proposition 1 to obtain agents' actual migration decision in $t$, which moves the economy to $t+1$. Repeating this process in $t+1$ delivers the agents' decision in $t+1$, which gives the allocation at the beginning of $t+2$. Iterative on this process forward until $T$ or until when the agents have perfect foresight in $T' < T$ delivers the evolution of the economy under evolving beliefs.

As an illustration, consider a model economy in which the productivity in one location is growing. Agents in the economy form beliefs about this productivity growth and make migration decisions accordingly, but their beliefs are overly optimistic. In each period, we first solve for the perfect foresight path; we then approximate around that perfect foresight for agents' actual decisions, which shifts the economy to $t+1$. Figure 2 illustrates how such an adjustment process can unfold. The vertical axis of the figure is the share of people in the location experiencing productivity growth; the dotted lines plot the shares according to agents' belief at each point in time; the solid line is the actual outcome; the dashed lines plot the outcome if agents gain perfect foresight at different periods. As the figure illustrates, in each period, a fraction of agents move to location 1, anticipating (mistakenly) that as the productivity in location 1 continues to grow, in the future, more people will move in. Every period, agents are surprised by the realization of productivity and revise their beliefs. This process continues until the last period of the economy. Throughout these periods, in addition to productivity, we can allow for other fundamentals of the economy, such as trade and migration costs, and regional amenities to vary over time.

In summary, Proposition 1 allows us to obtain a first-order solution to the model with stochastic environments on any path of productivity $z^T$. In the subsequent sections, we show that this approach generalizes tractably to accommodate uncertainty and heterogeneous beliefs. We now explain the rationale for choosing the deterministic path as the approximation point and discuss how to construct such a path in a stochastic environment for both ex-ante and ex-post analysis.

**The choice of the approximation point.** Our solution method uses a deterministic transition path with perfect foresight for approximation. In principle, any such path can be used. For example, in cases where the model has a steady state, one can use the steady state as an approximation point, as in the business cycle literature. Compared to this alternative, using the deterministic transition path has three advantages.

First, it is versatile. Our approach does not rely on the existence of a steady state, so it applies to non-stationary models. The model can be non-stationary if, for example, it has a finite horizon, or it has an infinite horizon but the fundamentals—or agents' belief about the fundamentals—follow a non-stationary process.

Second, it can accommodate the fact that in many applications of dynamic spatial economies, due to either the shock under investigation being large or other forces not directly related to the shock (e.g., the current distribution of workers being far away from the steady state), the economy might start from a point with drastically different economic landscapes from the steady state to which it converges. Using the steady state to approximate the decision today can lead to substan-

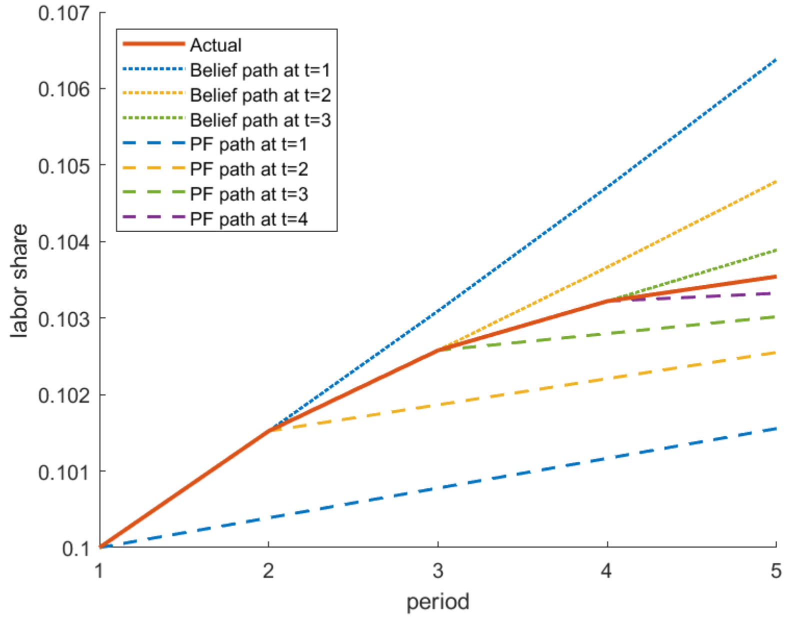



Figure 2: An Economy under Evolving Beliefs

Notes: simulations of an economy in which a location receives positive productivity shocks and agents in the model are too optimistic about the shocks. Plotted in the figure is the share of the population in the location receiving positive productivity shocks.

tial approximation errors. One might be tempted to think that such approximation errors are of second order so they do not matter materially for first-order approximations. To understand the nature of such errors, it is important to be precise about in what sense are the errors 'second order.' For concreteness, consider the approximation of a smooth single-variable function $f(x)$. Suppose we are interested in finding out $f(x_2+\delta) - f(x_2)$, i.e., the impact of a change of $\delta$ on $f(x)$ around $x_2$. First-order approximation around $x_2$ gives us $f(x_2+\delta) - f(x_2) = f'(x_2)\cdot\delta + O(\delta^2)$, with the error being second order in $\delta$, the shock that we are interested in study. If we approximate around $x_1$ instead, we would have

$$\begin{aligned} f(x_2+\delta) - f(x_2) &= f'(x_2)\cdot\delta + O(\delta^2) \\ &= f'(x_1)\cdot\delta + [f'(x_2) - f'(x_1)]\cdot\delta + O(\delta^2) \\ &= f'(x_1)\cdot\delta + f''(x_1)\cdot(x_2-x_1)\cdot\delta + O\left((x_2-x_1)^2\cdot\delta\right) + O(\delta^2). \end{aligned}$$

Thus, the additional error from approximating around $x_1$ is $f''(x_1)\cdot(x_2-x_1)\cdot\delta + O((x_2-x_1)^2\cdot\delta)$. Following this analogy, if the steady state ($x_1$) is far away from the actual path ($x_2$), then the approximation error is in the order of $(x_2-x_1)\delta$. Unless $x_2$ deviates only from steady states due to the shock $\delta$, in which case $x_2 - x_1$ is first-order in $\delta$ and hence $(x_2-x_1)\delta$ is second order in $\delta$,[7] the error is generally first order in $\delta$.

[7]This could be the case in, for example, the studies of business cycles, in which all deviations from the steady state are driven by shocks. In the application of dynamic spatial models, however, the economy can differ from the steady state due to other time-varying fundamentals or simply because the starting point of the economy is far from the steady state. In such scenarios, $x_2 - x_1$ can be much larger than $\delta$.

Third, approximating around the perfect foresight path can accommodate time variations in fundamentals other than productivities (e.g. migration costs, trade costs). Intuitively, our approximation can be interpreted as a cross-economy comparison, one between a stochastic economy with a given beliefs structure relative to a deterministic one. All fundamentals that are not the focus of a counterfactual would be canceled in this cross-comparison—in a sense, this approach is a stochastic version of the 'dynamic hat algebra.' As such, researchers do not need to recover all fundamentals that are not the focus of the counterfactual question. In contrast, to approximate the economy around a steady state, researchers first have to compuete the steady state and then recover the deviations of *all* fundamentals from their steady-state values. This feature will be especially convenient in ex-post studies when such time variations in other fundamentals are necessary to fit the data.

**Ex-ante versus Ex-post Analyses.** It is important to make a distinction between ex-ante and ex-post counterfactuals. In ex-ante counterfactuals, the objective of the analysis is in recovering how the economy evolves under a particular realization of productivity $z^T$, which might differ from the expected paths of $z$ implied by either the agents' beliefs or the true stochastic process. For decisions in each period, our approach requires first calculating a deterministic transition path with perfect foresight starting from that period, but it places little restriction on what perfect foresight paths to use. One example is to simply use the perfect foresight path calculated using $z^T$ (the dashed lines in Figure 2). Another example is to use the perfect foresight path corresponding to agents' expected future fundamentals (the dotted line in Figure 2). In the latter case, certainty equivalence implies that up to the first order, agents' decisions in $t$ under *that* perfect foresight path starting from $t$ coincide with their actual decision in $t$. As a result, we no longer need to use Proposition 1 to solve for the evolution of the economy.[8] Even in this case, Proposition 1 is still important as it provides the foundation for ex-post analysis, as we discuss in Section 2.5, and for incorporating uncertainty and heterogeneity beliefs, as we discuss in Sections 2.6 and 2.7.

In ex-post counterfactuals, researchers observed the evolution of the economy under a particular realization of the fundamentals and aim to recover the evolution of a counterfactual economy with a different path of fundamentals. Since the actual evolution are the result of both realized and unrealized but anticipated fundamentals, it is not the same as the perfect foresight path corresponding to the actual realization, and Proposition 1 cannot be readily used for counterfactual analysis. In the next subsection, we construct an algorithm to recover the perfect foresight path from the observed allocation and conduct counterfactuals.

### 2.5 Recovering Perfect Foresight and Counterfactual Paths in Ex-Post Studies

Following the previous discussion, in ex-post studies, researchers do not observe the perfect foresight path corresponding to the realized fundamentals. What is observed, instead, is the outcome of agents' decisions under evolving beliefs that do not necessarily coincide with the actual realization of fundamentals (in the example illustrated in Figure 2, the red solid line). In this subsection, we construct an algorithm to disentangle the role of beliefs from that of other time-varying fun-

[8]Indeed, in Figure 2, the dotted lines from period $t$ coincide with the actual evolution in the outcomes of period $t$.

damentals, recovering both the belief paths and the perfect foresight paths. We also show that for measuring the welfare of agents, the allocations in the actual and belief paths are sufficient and information on fundamentals such as migration costs is not needed. This result will prove useful for welfare evaluation in environments where agents do not have perfect foresight.

Suppose first that the researcher observed the actual allocation throughout all periods. Our algorithm proceeds in two steps. In the first step, we recover agents' beliefs about future paths at each point in time using a backward recursive algorithm that combines two insights. First, the observed allocation contains all information on trade costs and realized productivity, so we can use these outcomes in each period to solve for counterfactual *static trade equilibria* with different productivity beliefs or labor allocations as in Dekle et al. (2007). Second, agents' actual migration decisions are made to maximize their utility given their beliefs about future fundamentals, including migration costs, so we can use the actual migration decision recursively to recover *belief* paths without the need to recover time-varying migration costs.

Let the actual outcomes in period $t$ be denoted using a tilde, i.e., $\tilde{v}_{nt}$, $\tilde{w}_{nt}$, $\tilde{P}_{nt}$, $\tilde{\lambda}_{nit}$, $\tilde{\mu}_{nit}$, $\tilde{l}_{nt}$.

(i) Starting from period $T$, solve for the expected outcomes of the *static trade equilibrium* in period $T$ according to the beliefs of agents in period $T-1$, namely $\mathbb{E}(w_T|z^{T-1})$, $\mathbb{E}(P_T|z^{T-1})$ and $\mathbb{E}(\lambda_T|z^{T-1})$, by approximations around the actual outcomes in period $T$. The input for this step is the actual outcome in $T$ and the deviations in agents' belief in period $T-1$ about the productivity in $T$ from the actual productivity, namely $\mathbb{E}(\hat{z}_T|z^{T-1})$.

(ii) Append the output from the first step by $\tilde{\mu}_{niT-1}$, namely agents' *actual* migration decision in period $T-1$. As $\tilde{\mu}_{niT-1}$ is decided according to agents' belief in period $T-1$, together with the output from (i), it constitutes the solution to the agents' problem in period $T-1$ that is defined in Proposition 1.[9]

(iii) Solve for the expected outcomes in periods $\{T-1, T\}$ according to the beliefs of agents in period $T-2$ by approximating around the solution to agents' problem in period $T-1$, obtained from (ii). These outcomes include $\mathbb{E}(w_{nt}|z^{T-2})$, $\mathbb{E}(P_t|z^{T-2})$, $\mathbb{E}(l_{nt}|z^{T-2})$, $\mathbb{E}(\lambda_{nit}|z^{T-2})$, $\mathbb{E}(\mu_{nit}|z^{T-2})$, and $\mathbb{E}(v_{nt}|z^{T-2})$ for $t \in \{T-1, T\}$. The input to this approximation is the output from (ii) and the deviations in agents' belief in period $T-2$ about the productivity in $\{T-1, T\}$ from their belief in period $T-1$ about it, namely $\mathbb{E}(\log(z_t)|z^{T-2}) - \mathbb{E}(\log(z_t)|z^{T-1})$ for $t \in \{T-1, T\}$.

(iv) Append the output from (iii) with $\tilde{\mu}_{niT-2}$, agents' actual migration decision in period $T-2$. Together, $\tilde{\mu}_{niT-2}$ and the output from (iii) constitute the solution to the agents' problem in period T-2.

(v) Repeat (iii) and (iv) recursively backward until the first period is reached.

[9] The output of step (i) and $\tilde{\mu}_{niT-1}$ are both in level. However, the deviations of these level variables from the perfect foresight path from $T-1$ satisfy the conditions in Proposition 1. It is in this sense that they are the solution to the problem defined in Proposition 1.

This recursive algorithm recovers the solution to the agents' problem in each period and delivers expected values for all future outcomes according to agents' beliefs about productivity—in the example illustrated in Figure 2, these outcomes are represented using dotted lines.

In the second step, we recover the perfect foresight path from each period by approximating around the belief path from the same period—an application of Proposition 1. We can then use the perfect foresight path as the approximation point for counterfactual analysis as described in the previous subsection.

Importantly, in every step along the process, whenever we solve for a path as deviations from another path, the deviations are always defined between variables of the two paths in the same period, so time-varying trade and migration costs are canceled out in each step, which can be seen as a stochastic version of the 'dynamic hat algebra' approach developed in Caliendo et al. (2019).

**Extensions.** In illustrating the intuition, we have so far assumed that the model is finite and that the researcher observed the actual allocation until the end of the model. Now we discuss how the method can be generalized when these assumptions are relaxed.

Suppose the researcher observed all allocations up to period $T' < T$ but not between $T'$ and $T$, in which $T$ is either a finite number or infinity. Before the above algorithm can be applied, we first need to construct the path between $T'$ and $T$ according to the beliefs of agents in period $T'$, imposing additional assumptions about model fundamentals over $T'$ to $T$. For example, we can assume that all *other* fundamentals are either constant or change in a specific way after $T'$, and that $z$ still follows a stochastic process over which agents form beliefs.[10] With these assumptions, we can solve the agents' belief path based on the observed allocation in period $T'$, using a time-difference version of the model. The intuition is that the deviations in decisions and equilibrium outcomes in each future period from the observed allocation in period $T'$ can be written as a system of equations with the deviations in fundamentals between period $T'$ and future periods as the input.

Once we have recovered the belief path from $T'$, we can then start with step (iii) of the algorithm and iterate backward to recover all belief paths between 1 and $T'$. From these paths, we can use Proposition 1 to obtain the perfect foresight path starting from any period, or any other counterfactual paths.

In addition to the observed allocation, a key input into this exercise is agents' beliefs about fundamental productivity. In some applications, such data might not be readily available. If the researcher can gather some information about agents' beliefs about other endogenous outcomes that are affected by fundamental productivity, then we can proceed in a nested approach—by choosing the belief on fundamentals that leads to the best fit between the recovered belief path and agents' actual beliefs about endogenous outcomes. We will illustrate this approach in our second application.

**Welfare evaluation.** In applications of dynamic spatial models, researchers are interested in not only how allocations are affected by a shock, but also how agents' average welfare, measured using the value of locations $v_{nt}$, is affected. In environments with perfect foresight, such

[10]Some versions of such assumptions are unavoidable as researchers do not observe any data after $T'$.

calculation is straightforward for both ex-ante and ex-post analyses. In particular, in ex-ante analyses where researchers have made an assumption about fundamentals, $v_{nt}$ can be readily calculated recursively using the deterministic version of equation (2). In ex-post analyses where the researchers observe the perfect foresight transition path, the real wages of locations and migration rates along the perfect foresight path can be used to evaluate welfare. Intuitively, conditional on local wages, migration choice reflects the attractiveness of other locations, accounting for migration costs. Thus, welfare can be evaluated without the knowledge of underlying fundamentals.

In our environment with evolving beliefs, given any path of fundamental productivity, there are at least three notions of welfare (location values).[11] First, the expected values according to agents' beliefs at each point in time—in the example illustrated in Figure 2, the allocation associated with these values are denoted by the dotted lines. Second, the values if agents had perfect foresight at each point in time, with the corresponding allocation denoted by the dashed lines in Figure 2. Last but not least, the values that correspond to the actual allocation (hereafter the actual values)—one where agents make decisions based on their beliefs, which are subsequently revised as new data arrive.

The first two notions of welfare can be calculated analogously as in perfect foresight environments. For example, by evaluating the expected linearized version of equation (2) recursively backward using the expected outcomes according to agents' belief in $t = 1$, we can calculate the first notion of welfare in $t = 1$. For the third notion of welfare, which is arguably the most relevant notion in many applications of dynamic spatial models, however, such an approach does not work. The intuition is that agents' decisions on the actual path are shaped by beliefs that are subsequently revised, so the *actual* values in $t$ and $t+1$ are not directly linked through a forward equation in the form of (2). In the rest of this subsection, we show that the third notion of welfare can be written as the sum of 1) the first notion of welfare and 2) an adjustment due to the discrepancy between the actual fundamental path and the expectation, and that both items can be recovered from allocations without the knowledge of fundamentals.

Formally, let $\tilde{v}_{nt}$ be the *actual* value of location $n$ in period $t$, and let $v_{nt}^{t+1}$ be agents' expectation in period $t$ about the value of location $n$ in period $t+1$.[12] We have:

$$\begin{aligned}
\tilde{v}_{nt} &= U(\tilde{c}_{nt}) + \mathbb{E}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta\tilde{v}_{jt+1} - m_{njt} + \epsilon_j] \qquad (10)\\
&= U(\tilde{c}_{nt}) + \mathbb{E}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta v_{jt}^{t+1} - m_{njt} + \epsilon_j] + \mathbb{E}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta\tilde{v}_{jt+1} - \beta v_{jt}^{t+1}]\\
&= U(\tilde{c}_{nt}) + [\beta v_{nt}^{t+1} - \nu\log(\tilde{\mu}_{nnt})] + \sum_j \tilde{\mu}_{njt}[\beta\tilde{v}_{jt+1} - \beta v_{jt}^{t+1}].
\end{aligned}$$

The first equality says that the actual value of a location is the sum of the flow utility and a continuation value, with the latter depending on agents' actual choice (a function of $\epsilon_j$, which is summarized by the indicator function) and the expectation taken over the realizations of $\epsilon_j$. The

[11] The path of fundamentals could be either a hypothetical path or (in ex-post analyses) the realized path.

[12] Since in this exercise we are evaluating the welfare for a given path of fundamental, we ignore $z^t$ as an argument for all variables below.

second equality split the continuation value into two components linearly. The third equality exploits the idea that the actual migration decision is made according to the *expected* values, which implies $\mathbb{E}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta v_{jt}^{t+1} - m_{njt} + \epsilon_j] = [\beta v_{nt}^{t+1} - \nu \log(\tilde{\mu}_{nnt})]$.

Equation (10) thus writes the actual values in period $t$ recursively as a function of the observed allocation, the expected values in $t+1$ according to the belief in $t$, and the actual values in period $t+1$. Starting from the last period of the model, or the period around which agents gain perfect foresight, we can iterate on equation (10) backward to obtain the actual value associated with each location in any period.

We summarize the results in this section in the following proposition.

**Proposition 2.** *i The actual allocations and agents' beliefs about the fundamentals in each period are sufficient to recover the expected paths according to agents' beliefs and the perfect foresight paths.*

*ii The actual allocation and the belief path at each period are sufficient to recover the utility of agents. Information on other fundamentals is not needed.*

*Proof.* See Appendix A.2. □

Finally, it should be clear from the discussion in this section that the general principles of our solution method for both ex-ante and ex-post analysis rely on approximating a stochastic system around a deterministic path, using allocation on the deterministic path and the processes on the fundamentals as input. This approach can therefore accommodate many extensions, including multiple sectors, intermediate goods with input-output linkages, and other sources of dynamics. We will use an extended model with multiple sectors with intermediate goods in quantitative applications.

### 2.6 Incorporating Uncertainty

In this section, we incorporate the uncertainty in agents' perceptions of the future into their forward-looking decisions. As it might be intuitive, doing so requires extending our methodology to accommodate a second-order approximation of the stochastic spatial equilibrium, which we now proceed to do.

As before, we denote by $\bar{x}, \bar{y}$ variables in the deterministic environment with perfect foresight, namely $\bar{x}, \bar{y} \in \{\bar{w}_{nt}, \bar{P}_{nt}, \bar{v}_{nt}, \bar{z}_{nt}, ...\}_{n=1}^{N}$ and by $\hat{x}, \hat{y}$ the corresponding stochastic deviation from the deterministic equilibrium, that is $\hat{x}, \hat{y} \in \left\{\hat{w}_{nt}, \hat{P}_{nt}, \hat{v}_{nt}, \hat{z}_{nt}, ...\right\}_{n=1}^{N}$. To fix ideas, and following the same example as in the previous section, we derive the second-order approximation of the expected values in equation (8) as

$$\mathbb{E}_1 \hat{v}_{it+1} \approx \frac{\partial F_i}{\partial \bar{c}_{it+1}} \mathbb{E}_1 \hat{c}_{it+1} + \frac{\partial F_i}{\partial \bar{v}_{t+2}} \cdot \hat{v}_{t+2} + \frac{1}{2} \sum_{x,y} \frac{\partial F_i^2}{\partial \bar{x} \partial \bar{y}} \mathbb{E}_1 \hat{x}\hat{y}. \tag{11}$$

Equation (11) shows that moving to second-order approximation requires adding covariance terms between all arguments of $F_i$ to the first-order approximation given in equation (9). The coefficient for covariance $\mathbb{E}_1 \hat{x}\hat{y}$ is the cross-derivatives of $F_i$ with respect to $x$ and $y$ evaluated at the values on the deterministic path, which we denote by $\frac{\partial F_i^2}{\partial \bar{x} \partial \bar{y}}$. In Appendix A.3, we show that $\frac{\partial F_i^2}{\partial \bar{x} \partial \bar{y}}$ can be characterized as closed-form functions of the outcomes on the deterministic path (e.g., trade and migration shares). We also provide similar expressions for all equilibrium conditions in the stochastic economy, thereby establishing the following proposition:

**Proposition 3.** *Given the labor allocations in period $t$, agents' expectation at time $t$ about the deviations in exogenous fundamentals and endogenous variables at time $t' > t$ from the deterministic path with perfect foresight $\left\{ \mathbb{E}_t \hat{z}_{it'}, \mathbb{E}_t \hat{v}_{it'}, \mathbb{E}_t \hat{\mu}_{nit'}, \mathbb{E}_t \hat{l}_{it'}, \mathbb{E}_t \hat{w}_{it'}, \mathbb{E}_t \hat{\lambda}_{nit'}, \mathbb{E}_t \hat{P}_{it'} \right\}_{i=1,n=1,t'=t}^{N,N,T}$ and the covariance of these deviations $\{ e.g, \mathbb{E}_t \hat{z}_{it'} \hat{z}_{nt'}, \mathbb{E}_t \hat{v}_{it'} \hat{v}_{nt'}, \mathbb{E}_t \hat{w}_{it'} \hat{w}_{nt'}, \mathbb{E}_t \hat{P}_{it'} \hat{P}_{nt'}, \mathbb{E}_t \hat{\lambda}_{nit'} \hat{\lambda}_{nit'}, \mathbb{E}_t \hat{w}_{it'} \hat{P}_{nt'} \}$ solve a system of linear equilibrium conditions (in the order of $N^2 \times (T - t)$), with the coefficients of the terms being closed-form functions of the outcomes from the deterministic path.*

*Proof.* See Appendix A.3. □

Proposition 3 shows that the same system of equations as in Proposition 1, appended with second-order terms, characterizes the solution with second-order accuracy. However, while in the case of Proposition 1 we can simply invert the system of equations for solutions, here the system is under-ranked—there are far more second-order terms in these equations than the number of equations. To obtain the solution with second-order precision, we adopt a two-step procedure.

In the first step, we calculate the covariance terms $\mathbb{E}_1 \hat{x}_{t'} \hat{y}_{t'}$ based on the simulations of the first-order solutions obtained using Proposition 1. Specifically, we simulate $S$ paths of fundamental productivities. For each simulated path $s = 1, 2, ..., S$, we use Proposition 1 recursively to obtain the outcomes of the economy, denoted by $\hat{x}_{t'}(s)$ or $\hat{y}_{t'}(s)$. We then approximate $\mathbb{E}_1 \hat{x}_{t'} \hat{y}_{t'} \approx \frac{1}{S} \sum_{s=1}^{S} \hat{x}_{t'}(s) \hat{y}_{t'}(s)$. In the second step, we then solve the linear conditions in Proposition 3 for the remaining terms, replacing the covariance terms with these simulated values. The resulting solution will have second-order accuracy. To give some intuition of this result, notice that

$$\begin{aligned}
\mathbb{E}_1 \hat{x}_{t'} \hat{y}_{t'} &= \frac{1}{S} \sum_{s=1}^{S} \left[ \hat{x}_{t'}(s) + o\big(\hat{x}_{t'}(s)\big) \right] \cdot \left[ \hat{y}_{t'}(s) + o\big(\hat{x}_{t'}(s)\big) \right] + O\left(1/\sqrt{S}\right) \\
&= \frac{1}{S} \sum_{s=1}^{S} \hat{x}_{t'}(s) \hat{y}_{t'}(s) + \underbrace{o\big(\hat{x}_{t'}(s)\big) \cdot o\big(\hat{y}_{t'}(s)\big) + \hat{y}_{t'}(s) \cdot o\big(\hat{x}_{t'}(s)\big) + \hat{x}_{t'}(s) \cdot o\big(\hat{y}_{t'}(s)\big)}_{\text{third-order terms}} + O\left(1/\sqrt{S}\right)
\end{aligned}$$

where $o\big(\hat{x}_{t'}(s)\big)$ and $o\big(\hat{y}_{t'}(s)\big)$ represent the second-order errors in solving for $\hat{x}_{t'}(s)$ and $\hat{y}_{t'}(s)$ using Proposition 1, $O(\frac{1}{\sqrt{S}})$ denotes the simulation error for covariance. For large enough $S$, the approximation error due to simulation, $O(\frac{1}{\sqrt{S}})$, is arbitrarily small. The only remaining source of error is $o\big(\hat{x}_{t'}(s)\big)$ and $o\big(\hat{y}_{t'}(s)\big)$. Both errors are of second order, and they enter the approximation in product with other first-order or second-order terms. As a result, the errors in this approximation are of the third order, and by solving the system of linear equations for the remaining variables, our solution has second-order precision. We can then use Proposition 3 recursively to obtain the outcome of the economy under any given path of fundamentals with second-order precision that is needed to speak to the effect of uncertainty.

Our second-order approximation relates to the perturbation method used in the study of business cycles in macroeconomics (see Kim et al., 2008; Schmitt-Grohé and Uribe, 2004 and the references thereto), with two main differences. First, while the typical perturbation method focuses on a state-space representation of the economy, our method focuses on the sequence space. As discussed previously, this difference enables us to easily accommodate non-stationary models, allowing fundamentals to vary arbitrarily and agents' beliefs about these fundamentals to depend on history in a flexible way.[13] Second, unlike in conventional applications of perturbation methods, where the researchers need to calculate numerically the (first and second-order) derivatives of policy functions with respect to the state, we take advantage of the fact that in a widely used class of dynamic spatial models, the derivatives necessary for second-order approximation can be derived as closed-form functions of the outcomes on the approximation path. This enables us to accommodate a large number of locations and location-specific stochastic fundamental processes, which could be important in a large class of models.

### 2.7 Incorporating Heterogeneous Beliefs

Our analyses so far have focused on the case where all agents share a common belief $f$. In environments with stochastic fundamentals, agents might disagree about their expected future path of fundamentals and make different decisions, which impacts general equilibrium outcomes. Moreover, in such environments, agents' decision must take into account not only their own beliefs but also the beliefs of other agents as their decision also affects equilibrium outcomes. In this subsection, we extend our framework and methodology to a setting with heterogeneous beliefs.

Without loss of generality, assume that there are two different beliefs about the future path of fundamental productivities, each held by one group of agents.[14] We denote the two groups of agents by $A$ and $B$, with their beliefs denoted by $f^g(z_{t+1}|z^t)$, $g \in \{A, B\}$. We assume that the existence of these heterogeneous beliefs is common knowledge. In other words, agents are aware of the existence of the other group with a different belief but think that their own belief is the correct one. As all decisions of all agents interact in equilibrium, each group makes decisions considering the decisions of the other groups guided by their own belief. For instance, group $A$ agents might believe that the future productivity will be higher in a given location and move to that location. Their decision depends on the belief of group $B$. If $B$ thinks differently and moves less into that location, then group $A$ might have a stronger incentive to move into the location to take advantage of the lack of competition in the labor market there.

We now define the stochastic sequential equilibrium in this environment.

**Definition 2.** ***Stochastic sequential equilibrium with heterogeneous beliefs****. A stochastic sequential equilibrium with heterogeneous beliefs is a set of state-contingent prices $\left\{w_{nt}(z^t), P_{nt}(z^t)\right\}_{n=1,t=1}^{N,\infty}$, location values $\left\{v_{nt}^g(z^t)\right\}_{n=1,t=1,g\epsilon\{A,B\}}^{N,\infty}$, and allocations $\left\{\mu_{nit}^g(z^t), \lambda_{nit}(z^t), l_{nt}^g(z^{t-1})\right\}_{n=1,i=1,t=1,g\epsilon\{A,B\}}^{N,N,\infty}$ such*

[13]The number of simulations required for accuracy might be large if the processes have long history dependence, but simulations are easily parallelized.

[14]Our approach generalizes to any finite number of belief types, although the computational burden increases with the types of beliefs. Even with just two belief types, through the composition of agents' belief types, our model can accommodate variations in beliefs across locations.

*that $w_t(z^t), P_t(z^t), \lambda_{nit}(z^t)$ solve for the trade equilibrium given by the bilateral trade shares (5), the local prices (6), and the labor market clearing condition (7), and $v^g_{nt}(z^t), \mu^g_{nit}(z^t), l^g_t(z^{t-1})$ solve for the dynamic migration decisions given by value function (2), the gross flows equation (3), the law of motion of labor (4), for each group $g$, and consistent with beliefs $f^g(z_{t+1}|z^t)$.*

Under Definition 2, agents correctly anticipate the decisions of all other agents in all scenarios $z^t$, despite their disagreement on the likelihood of $z^t$. In this sense, this definition corresponds to an equilibrium of full rationality. Solving this equilibrium fully is clearly infeasible for the same reason that solving the equilibrium with homogeneous beliefs is infeasible. In fact, with heterogeneous beliefs, even the local approximation solution method established in Proposition 1 involves a curse of dimensionality challenge that is akin to one in the literature of higher-order beliefs (Townsend, 1983). In what follows, we first discuss the nature of this challenge and then make progress to gain tractability by proposing an approximate solution.

Suppose we follow the approach in Proposition 1. Consider the problem of group $A$, which takes as given the migration decision of group $B$, $\mu^B_{nit}(z^t)$. As discussed before, forward-looking decisions by group $A$ depend on the expected future path of real wages that are in part shaped by the forward-looking decisions by group $B$. Hence, solving for decisions by group $A$ in the first period requires considering the belief of group $A$ about group $B$ future decisions for all $t \geq 2$. Analogously, future decisions by group $B$ in turn depend on their belief about expected decisions by group $A$, introducing higher-order beliefs into the problem.

Figure 3a illustrates this dependence more concretely. In the example, in deciding what to do in the first period, group $A$ form expectation about the decision of group B in the second period and beyond, which can be denoted by $\mathbb{E}^A_1(\mu^B_{int})$, $t = 2, 3, ..., T$. In turn, $\mathbb{E}^A_1(\mu^B_{int}), t = 2, 3, ..., T$ depends on A's belief about B's belief at each $t$ about the future ($t'$) decision of $A$, namely $\mathbb{E}^A_1\mathbb{E}^B_t(\mu^A_{int'}), t' = t, 3, ..., T$. Following the same argument, $\mathbb{E}^A_1\mathbb{E}^B_t(\mu^A_{int'}), t' = t, ..., T$ depends on A's belief about B's belief about A's belief about B's decision in each future period $t'' > t'$, namely $\mathbb{E}^A_1\mathbb{E}^B_t\mathbb{E}^A_{t'}(\mu^B_{int''}), t'' = t', ..., T$. This process continues until the end of the model and implies that first-order approximation to A's problem in the first period requires us to solve higher-order beliefs about future decisions, which depend on higher-order beliefs about fundamental productivity (e.g., $\mathbb{E}^A_1\mathbb{E}^B_2\mathbb{E}^A_3 z_4 ...$). In the example illustrated in Figure 3a, one must solve for the outcome on each of the nodes, which is infeasible when $z_t$ takes continuous support.[15]

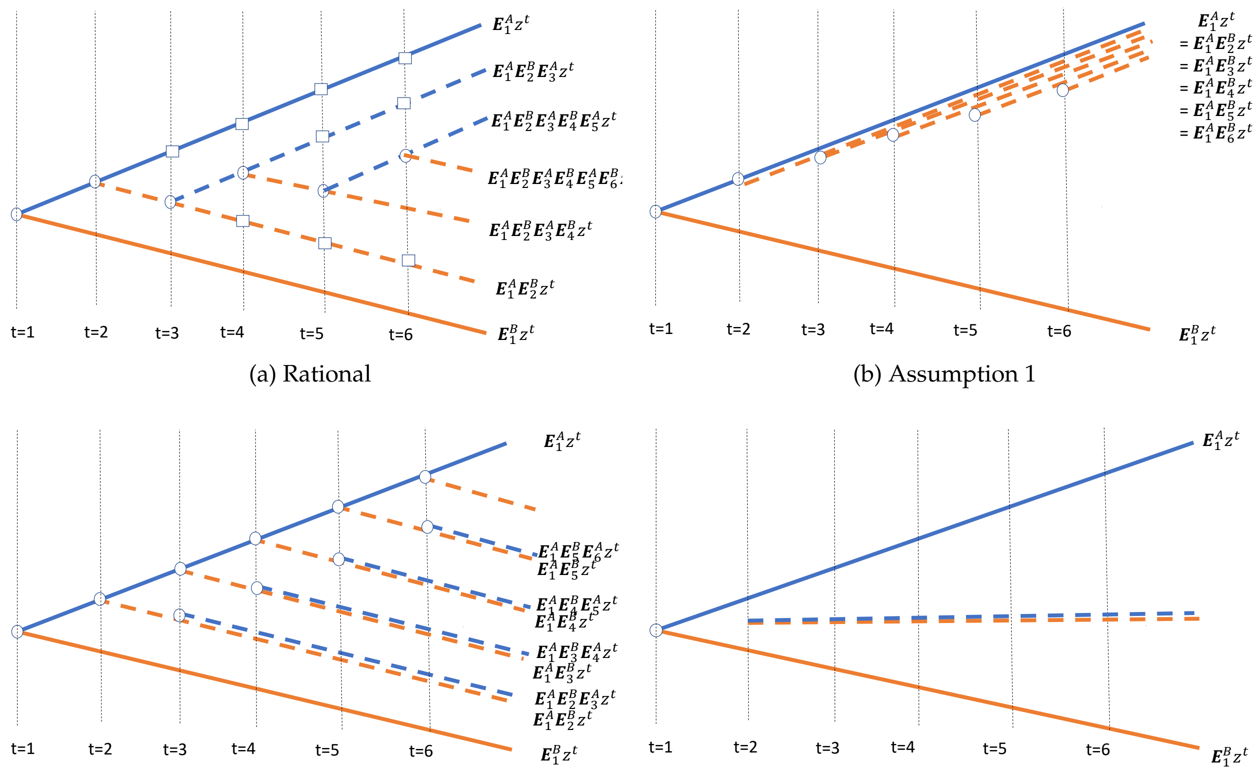
To retain tractability, we impose three alternative structures on the beliefs of agents.

**Assumption 1.** *At period t, each type thinks (wishfully) in t + 1 and beyond, the new data will vindicate their own belief and convince the other type.*

**Assumption 2.** *Agents understand that both types are stubborn and will stick with their beliefs, group A thinks group B thinks that A will be convinced by the data in the next period, and vice versa.*

**Assumption 3.** *Agents think their beliefs will converge in the next period.*

[15] In the special case where A and B share the same belief, these high-order beliefs collapse to a common expectation because of the law of iterative expectations. For example, $\mathbb{E}^A_1\mathbb{E}^B_2\mathbb{E}^A_3 z_4 = \mathbb{E}_1 z_4$.




(a) Rational

(b) Assumption 1

(c) Assumption 2

(d) Assumption 3

Figure 3: Structures of Heterogeneous Beliefs

Panels (b) (c) and (d) of Figure 3 illustrate the beliefs structure under these three assumptions. Under the first assumption, since each group expects that realized fundamental productivities will convince the other agent that their own beliefs are correct, only the second-order beliefs show up in the equilibrium definition. Under the second assumption, group A thinks that group B thinks group A will be convinced, and vice versa, so only up to third-order beliefs show up. Moving from Assumption 1 to Assumption 2, the complexity of the model increases. If we add more and more higher-order beliefs, the model will approach the rational stochastic equilibrium at the expense of computational complexity. Instead of going into higher-order beliefs, one can consider a setting (under Assumption 3) where both groups believe that they will converge to a common belief in the next period. These three assumptions capture a wide range of empirically relevant settings. Imposing any of these assumptions, we are able to retain tractability in the stochastic spatial framework with heterogeneous beliefs, as established in Proposition 4.

**Proposition 4.** *Under each of the three assumptions, the local solution to the model can be characterized by a system of linear equations and can be solved analogously to Proposition 1.*

*Proof.* See Appendix A.4. □

**Summary and extensions.** In this section, we have developed a generic stochastic dynamic spatial model with flexible beliefs. We derive an approximate solution for this model around a deterministic path with perfect foresight. We then extend this method to accommodate three empirically relevant settings: ex-post studies where researchers—in order to conduct counterfactuals—need to reconcile the observed allocations through evolving beliefs and/or other time-varying fundamentals; and uncertainty and heterogeneity in agents' beliefs. In the rest of this paper, we apply our methodology to shed light on the role of these departures from perfect foresight using two applications. To isolate the importance of each of them, we will focus on one at a time, but these departures can be combined, as we establish in Proposition 5.

**Proposition 5.** *Under assumptions 1-3, the stochastic spatial framework with heterogeneous beliefs can be solved with second-order accuracy as described in Section 2.6 so uncertainty can be incorporated. The results described in Sections 2.6 and 2.7 can also be applied in ex-post studies, where researchers observe outcomes shaped by the decisions of agents, whose beliefs are uncertain or heterogeneous and are interested in finding out the allocation corresponding to a counterfactual fundamental path in this environment.*

*Proof.* See Appendix A.5. □

In Appendix A.5, we note that in ex-post studies with heterogeneous agents, our approach differs depending on whether the migration flows by belief type are observed. There, we consider two polar cases: one in which the researcher can observe migration flows by belief type in all periods; one in which the researcher cannot observe migration flows by belief type but can observe total bilateral migration for each period and one cross-sectional distribution of people by their belief type. We show how to conduct counterfactuals in both cases and discuss how to proceed when the data available to the researcher falls between the two cases.

We now turn to quantitative analysis, in which we explore the roles of evolving, uncertain, and heterogeneous beliefs in dynamic spatial models, applying our methodologies to two applications—climate change and China's productivity shock.

# 3 Ex-ante Application: Climate Change

Our first application studies the effects of climate change on welfare and the spatial distribution of economic activities in the United States. As the full impact of climate change is largely yet to unfold, we use this example to illustrate an ex-ante application, focusing on two departures from perfect foresight that are especially relevant in this context: uncertainty and heterogeneous beliefs. In the first one, agents are uncertain about the path of the temperature rise. In the second one, agents have two different views of climate change—they either believe it is happening or are climate skeptics who do not think climate change will occur.

## 3.1 Setup and Data

We extend the baseline spatial model presented in the previous section to incorporate multiple sectors (agriculture, manufacturing, and non-tradable service industry), and intermediate goods. We denote sectors by $j, k \in \{A, M, S\}$. For example, $v_{nj}^t$ is the value associated with sector $j$ in location $n$ in period $t$; $\mu^t_{nj,ik}$ is the fraction of people moving from location $n$ sector $j$ to location $i$ sector $k$; $\lambda^t_{ni,j}$ is the share of location $n$'s expenditures in sector $j$ that is purchased from $i$.

With a focus on the United States, we map locations in the model to either the states in the United States or 28 other major countries and aggregated regions (see Appendix B for the list of countries and regions in included in the quantitative analysis). Migration can take place between U.S. states and sectors, but not internationally. In addition, as one of the focuses of our quantitative analysis is uncertainty, agents' risk attitude is important. Accordingly, we specify the flow utility from consumption as

$$U(c^t_{ni}) = \frac{(c^t_{ni})^{1-\sigma}}{1-\sigma},$$

where we set the risk-aversion parameter $\sigma = 3$. We set a migration elasticity of $1/\nu = 0.187$ following Caliendo et al. (2019) and a trade elasticity $\theta = 4.55$ from Caliendo and Parro (2015). We choose $\beta$ so the annual discount rate is 4%.

We study the role of uncertainty and heterogeneous beliefs in shaping the economic impact of the rising temperature over the period 2014-2100. As discussed in the previous section, doing so requires obtaining the initial allocation (trade, production, migration) for the year 2014 and computing a deterministic path with perfect foresight starting from that allocation. We will use a path of fundamentals that reflects the average effect of climate change and compute the corresponding deterministic path. After doing that, we first apply Proposition 3 to study the effect of uncertainty about the path of climate change, and then apply Proposition 4 to quantify the role of heterogeneous beliefs about climate change.

We start by describing the data used to obtain the initial allocation for the year 2014. Additional details are presented in Appendix B.

**Trade and production data.** We obtain bilateral trade flows across countries, gross output, value-added, and the input and final consumption shares across sectors and countries from the World Input-Output Database (WIOD).[16] Value added across individual U.S. states and sectors is obtained from the Bureau of Economic Analysis (BEA). We construct gross output across states and sectors using the sectoral U.S. value-added shares in gross output from WIOD.

We obtain bilateral trade between U.S. states and foreign countries in 2014 from the U.S. Census international trade data. We scale the imports and exports of individual states across sectors to match aggregate exports and imports between the United States and other countries in the WIOD.

Bilateral trade flows across U.S. states are obtained from the Commodity Flow Survey (CFS), which contains shipment information between and within states for twelve manufacturing sectors. The closest survey year to our sample period is 2012. For the manufacturing sector, we scale the shipment values in the 2012 survey so that the aggregate shipment values (local and interstate flows combined) are consistent with the total U.S. manufacturing output that is absorbed domestically in the WIOD for the year 2014. For the agricultural sector, we impute the interstate trade

[16] Consistent with our assumption that services are non-tradable, all production of services in WIOD are assumed to be used domestically.

flows from $i$ to $n$, $X^t_{in,A}$, using the following gravity specification:

$$\log(X^t_{in,A}) = \beta_1 \log(X^t_{n,A}) + \beta_2 \log(Y^t_{i,A}) + \beta_3 \log(\text{dist}_{in}) + \beta_4 \mathbb{I}(i = n),$$

where $X^t_{n,A}$ is the total absorption in the agricultural sector in location $n$, $Y^t_{i,A}$ is the gross production in agriculture in location $i$, $\text{dist}_{in}$ is the distance between the two locations, and $\mathbb{I}(i = n)$ is an indicator function that takes the value of one for the within-state flows. Gross output and total absorption in the agriculture sector across locations, $X^t_{n,A}$ and $Y^t_{i,A}$, are obtained using regional value added data and value added shares as we described above. We obtain estimates of the coefficient $\beta_s$ by running this gravity specification using CFS data on manufacturing shipments, and we then project the interstate trade flows for the agriculture sector using the estimated coefficients as well as the production data for the agriculture sector.

**Migration.** We construct annual gross mobility flows across U.S. states and sectors for the year 2014 from the job-to-job flows provided by the U.S. Census based on the Longitudinal Employer-Household Dynamics Database (LEHD), which counts the number of movements between states and sectors at a quarterly frequency. We calculate the number of stayers in each state-sector cell as the difference between total employment and outflows.

We now turn to construct a path fundamentals that reflects the impact of climate change. In particular, as described in what follows, we use temperature projection data to compute the impact of climate change on fundamental productivity over time.

**Future temperature rise and productivity damage.** Our temperature projection data are from the Climate Impact Lab, which provides temperature forecasts for world regions and the U.S. states under different emissions scenarios, namely RCP 4.5 and RCP 8.5 scenarios. Each scenario also contains three possible trajectories for temperature rise (median, 5th percentile, and 95th percentile of future temperature). In each case, the projection is reported for four periods: 1986-2005, 2020-2039, 2040-2059, and 2080-2099. We use the RCP 4.5 scenario, which is an intermediate scenario, and interpolate the temperature within these time frames linearly to obtain a temperature path (see Appendix B for additional details).

We specify the following damage function by which rising temperature affects regional productivity:

$$\begin{aligned} \log(\bar{z}^t_{nj}) &= \delta_{nj} + \alpha^1_j \bar{C}^t_n + \alpha^2_j (\bar{C}^t_n)^2 \\ \implies \Delta \log(\bar{z}^t_{nj}) &= \alpha^1_j \Delta \bar{C}^t_n + \alpha^2_j \Delta (\bar{C}^t_n)^2 \end{aligned} \tag{12}$$

where $\bar{z}^t_{nj}$ is the productivity in location $n$ industry $j$, as introduced before, $\delta_{nj}$ is a location-industry specific constant, $\bar{C}^t_n$ is the average annual temperature (in Celsius), and the coefficients $\alpha^1_j$ and $\alpha^2_j$ determine the quadratic impact of temperature on productivity. Given a path of temperature and estimates for $\alpha^1_j$ and $\alpha^2_j$, equation (12) gives us the change in productivity due to temperature rise.

We calibrate $\alpha^1_j$ and $\alpha^2_j$ based on the findings of Cruz (2023), who uses annual production data from world regions to estimate the long-run effect of climate change on sectoral productivity.[17]

[17] Cruz (2023) estimates damage functions for construction, trade and transportation, finance, and government sepa-

Table 1: The Quadratic Impact of Temperature on Productivity

| | Agriculture | Manufacturing | Service |
|---|---|---|---|
| $\alpha_j^1$ | 2.998 | 1.339 | 1.292 |
| | (1.268) | (1.344) | (1.051) |
| $\alpha_j^2$ | -0.143 | -0.0615 | -0.0485 |
| | (0.0368) | (0.0472) | (0.0405) |

Note: Estimates and standard errors (in parenthesis) based on Cruz (2023).

Table 1 reports the estimates. Although the coefficients differ across sectors, all three sectors show an inverted-U relationship between temperature and productivity. Importantly, these estimates are also quite uncertain, reflecting the inherent difficulty in inferring the economic tolls of climate change. We will incorporate the uncertainty in these estimates into agents' beliefs.

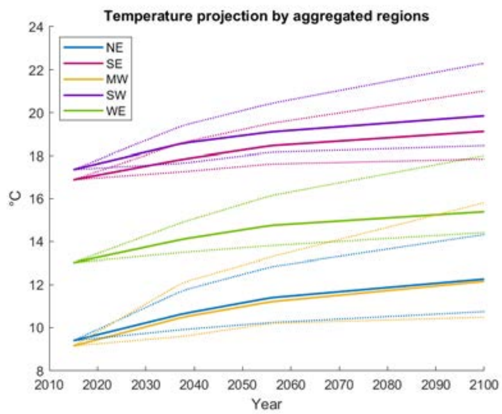
Panel (a) of Figure 4 reports the paths of temperature under the medium carbon emission scenario under RCP 4.5 for five broad U.S. regions: Northeast (NE), Southwest (SW), West (WE), Southeast (SE), and Midwest (MW). The solid lines present the median projection path for each region, which shows an across-the-board increase in temperature. The dotted lines in the figure show the 5% and 95% percentiles of the temperature path for each region. The wide bands in the figure reflect that even within the moderate carbon-emission scenario, the scientific community is fairly uncertain about how temperature will evolve. Also, we can see in the figure that the temperature is higher in the two southern regions (SW and SE) than in the rest of the country. Hence, given the inverted U relationship between temperature and productivity, even though the temperature rise is similar across the country, its negative impact will be larger in the South.

Panels (b), (c), and (d) of Figure 4 plot the weighted average of the change in $\log(z_{ni}^t)$ for the broad U.S. regions calculated according to (12). We can see in the figure that productivity decreases substantially in all three sectors in the South, but the decrease is particularly pronounced in agriculture, where $\log(z_{nj}^t)$ decreases by almost 300 log points.[18] On the other hand, the West sees only a modest decrease in productivity in all sectors; the Northwest and Midwest see an increase in productivity. Although not shown in the figure, similar patterns hold for other countries; in general, high-latitude countries, such as Canada and Russia, experience an increase in productivity from climate change, whereas productivity in low-latitude countries such as India is hurt.

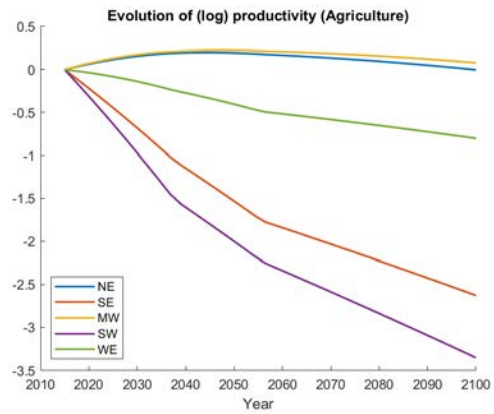
As the environment described above also makes it clear, in this application we take climate change (and its impact on the model fundamentals) as given, and we abstract from endogenous changes in carbon emission that are considered in dynamic integrated models of climate (e.g., Nordhaus, 1994). The purpose of this simplification is to focus on the methodology to study the role of uncertainty and heterogeneous beliefs. We note, however, that it is straightforward to

rately. We aggregate these coefficients based on their weight in the U.S. GDP to arrive at the coefficients for the broad service sector in our model. The productivity measure in Cruz (2023) is value-added productivity. We adjust these estimates by value-added shares to be consistent with our definition of $z_{nj}^t$ as the sectoral total factor productivity parameter.

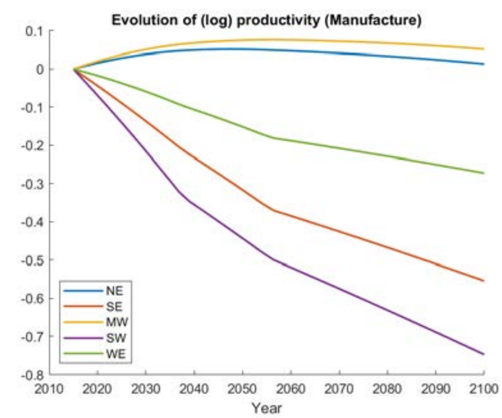
[18]Recall that $z_{nj}^t$ is the scale parameter of the Frechet distribution. In a closed economy, local TFP is $(z_{nj}^t)^{\frac{1}{\theta}}$. In that case, 300 log points decrease in $z$ maps into 66 log points decreases in TFP when $\theta = 4.55$.



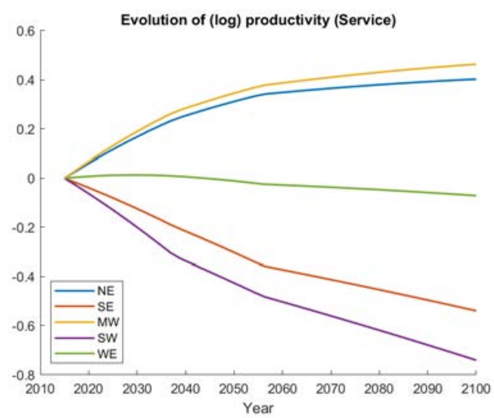

(a) Temperature projection




(b) Agricultural productivity




(c) Manufacturing productivity




(d) Service productivity

Figure 4: Temperature Projection and Productivity Path

Notes: Panel (a) shows the evolution of temperature for broad U.S. regions. The solid lines are the medium of the forecast, whereas the dotted lines are the 5th and 95th percentiles of the forecast. Panels (b), (c), and (d) are the evolution of productivity under the median temperature for broad U.S. regions.

extend our model to incorporate endogenous temperature rise.[19]

**Constructing a deterministic path with perfect foresight.** With the data on initial allocation, we first construct a deterministic path, around which we will apply Propositions 3 and 4. To do so, we need to take a stand on the paths of all fundamentals underlying the deterministic path. We assume that on this deterministic path, local productivity evolves according to the damage function (12) and the median temperature projection under RCP 4.5, as plotted in Figure 4.[20] We then obtain the deterministic path by solving a time-difference version of the model as in Caliendo et al. (2019), assuming that all other fundamentals (trade and migration costs) are constant after 2014. Intuitively, the initial allocation contains information on the level of model fundamentals in 2014; when combined with assumptions on the changes in these fundamentals along the deterministic path, it gives us all information necessary to solve for the model. In doing this time difference, we assume that the decision of agents in 2014 are made under correct expectations about future fundamentals. We make this assumption so our benchmark results on climate change with perfect foresight are comparable to the literature, all of which have imposed a perfect foresight assumption. However, it is worth noting that it is conceptually straightforward to allow the 2014 decision to be made under a different belief than the one implied by equation (12) and the median temperature projection.[21]

### 3.2 Climate Change versus No-Change Scenario

To establish a benchmark result against which we will evaluate the importance of uncertainty and heterogeneous beliefs, we first compare the deterministic path with productivity evolving according to equation (12), constructed in the previous subsection, to a counterfactual deterministic path in which productivity does not change after 2014—the no-change scenario. The difference between these two paths is the economic effect of perfectly anticipated climate change.

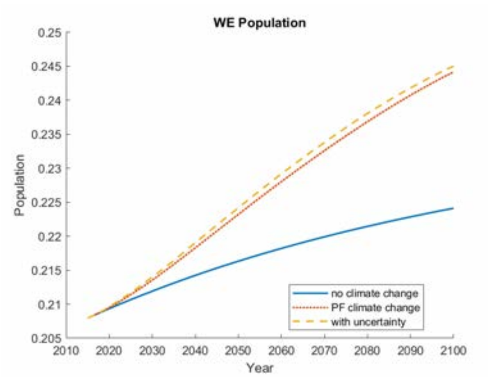
The red dotted lines in Figure 5 show the evolution of population across locations in the United States with perfectly anticipated climate change. As temperature rises, the population moves away from the South to the rest of the country. For example, by 2100, the two southern regions together account for only 27 percent of the U.S. population, down from 37 percent at the beginning of the period. The blue solid lines in the figure show the evolution of population across space under the 'no-change' scenario. We can see that if not for climate change, the population in the South location would increase, and the population in the other regions would decline.

Figure 6 plots the difference in welfare across individual states in the United States between the climate change scenario and the no-change scenario, with welfare measured as the change

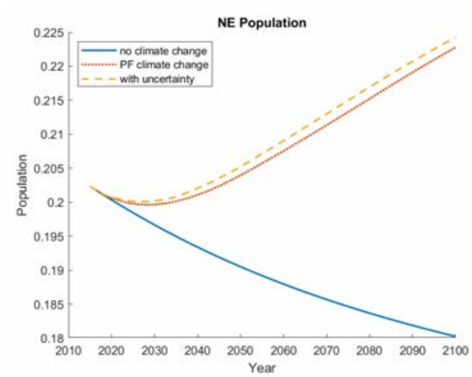
[19] In such extension, one can study the impact of uncertainty, evolving, and heterogeneous beliefs about how carbon emissions affect temperature and productivity by applying the propositions in the previous section.

[20] In principle, the deterministic transition path can be computed under any path of future temperature rise; in practice, the quality of the approximation is higher when the deterministic path used is closer to the ones used in subsequent analysis. For this reason, we will use the average value for the predicted temperature rise.

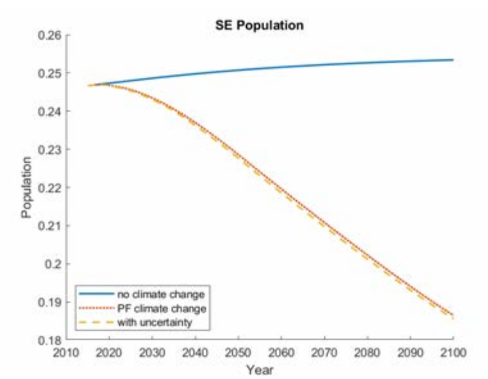
[21] In that case, we can first solve a time-difference version of the model to obtain the sequential equilibrium corresponding to agents' belief. We can then use Proposition 1 to obtain the deterministic path corresponding to fundamentals with median temperature change. In the same spirit, arbitrary time variations in trade and migration costs (and arbitrary beliefs about these) can also be incorporated.



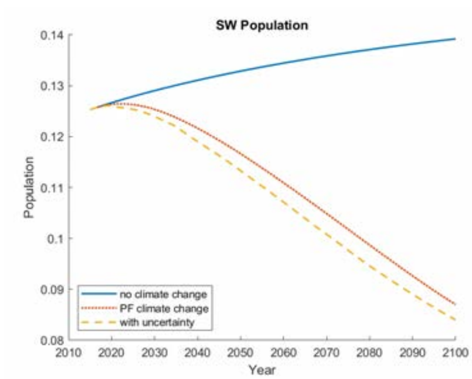

(a) West




(b) Northeast




(c) Southeast




(d) Southwest

Figure 5: Climate Change and Spatial Reallocation

Notes: The figure shows each region's share of the population over time under different scenarios.

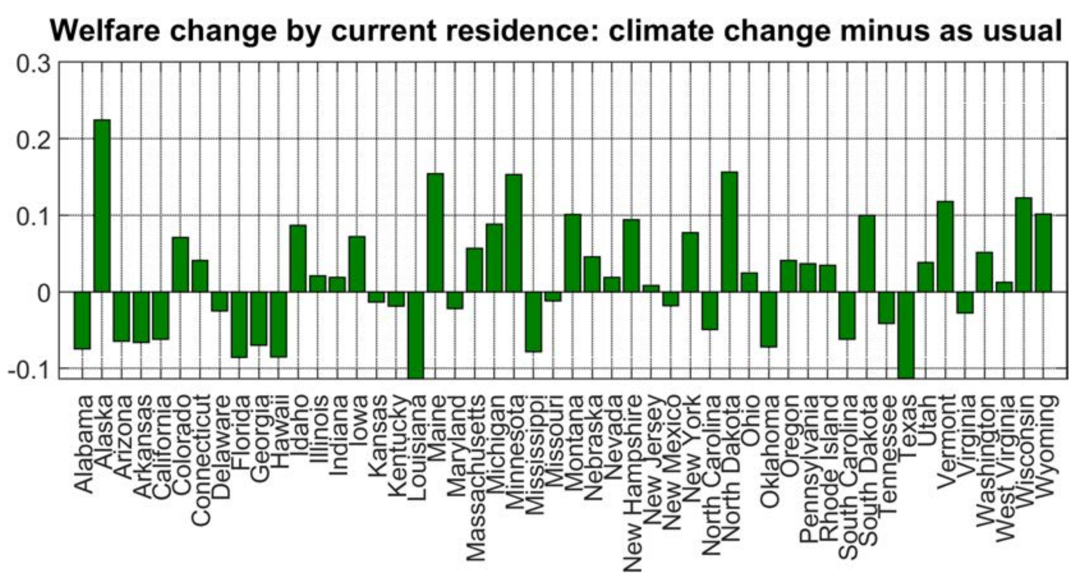



Figure 6: The Welfare Effect of Climate Change for U.S. States

Notes: This figure plots the percentage difference in the welfare of the current residents in the presence of climate change compared to in the absence of it.

in consumption equivalent.[22] Warmer temperatures due to climate change makes households located in northern states better off, with the gains the largest in currently colder states. On the other hand, southern states suffer welfare losses. For instance, households' welfare in Texas and Louisiana falls by more than 10%. Overall, the population-weighted average welfare loss for the United States as a whole is 0.9%, in line with recent estimates using deterministic models (e.g. Cruz and Rossi-Hansberg, 2023, Cruz, 2023).

### 3.3 The Role of Uncertainty

We now turn to study the dynamic effects of climate change under uncertainty. As we described before, Panel (a) of Figure 4 shows considerable uncertainty on the future path of temperature. Moreover, Table 1 suggests there is also uncertainty on how temperature affects productivity.

We apply Proposition 3 to study how such uncertainty affects the welfare and population reallocation within the United States. Specifically, for each period $t = 1, .., T$, we implement the second-order approximation around a perfect-foresight path by simulating $S$ paths of future temperature rise for U.S. states and other countries. We solve the sequential equilibrium under each of the simulations and then use the moments calculated from these $S$ solutions to solve for agents' decision in period $t$, which then shifts the economy into the next period.

Before describing our findings, we discuss three additional aspects of this exercise. First, we introduce uncertainty in temperature by adding mean-zero normal shocks around the median forecast for future temperatures. We calibrate the standard deviation of the shock to match the difference between the 5th and the 95th percentile of the future temperature distribution. Our assumption is essentially that the agents *today* have the same information and belief as the scientific community on future temperature rise. In addition to their beliefs today, to solve the model with uncertainty, we also need to specify agents' beliefs in future periods under different realizations. The challenge is that while the belief of the scientific community today can be measured, how such beliefs will be revised in the future as new data arrive is necessarily speculative. For example, if in the year 2030, the realization of temperature is higher than the median forecast made today for that year, will the future belief in 2030 be revised up or down? Will the perceived uncertainty for the future belief be higher or lower? To make progress, we assume that in any period $t$, agents' belief about the distribution for the period $t' \geq t$ is a normal distribution with mean $C_n^{t,t'}$ and variance $(\sigma_n^{t,t'})^2$. We further assume that $C_n^{t,t'}$ follow a mean-reverting process that converges to the median forecast of temperature in $t'$ by the scientific community today and that the uncertainty in period $t'$ temperature perceived by workers standing at $t \leq t'$ only depends on $t' - t$. Formally,

$$C_n^{t,t'+1} - \bar{C}_n^{t'+1} = 0.8(C_n^{t,t'} - \bar{C}_n^{t'}) \text{ and } \sigma_n^{t,t'} = \sigma_n^{1,t'-t+1}. \tag{13}$$

The term $\bar{C}_n^{t'}$ is the median forecast for period $t'$ temperature from $t = 1$. This setup implies that if the realization of $C_n^t$ is above $\bar{C}_n^t$, then agents think in the future $t'$, the average temperature will be higher than but eventually converge to $\bar{C}_n^{t'}$. We choose a persistent parameter of 0.8, which implies

[22] We aggregate across sectors in each location using population shares.

a moderate convergence speed. We calibrate $\sigma_n^{1,t'-t}$ to match the belief dispersion today described in Panel (a) of Figure 4 for each U.S. state.

Second, to capture the fact that climate change affects all states, in simulating future temperature paths, we assume that the percentiles of the draws for each state are perfectly correlated. For example, if Colorado has a draw for a period that is the 99th percentile according to its temperature distribution, then Utah also receives a 99th percentile draw from its distribution.

Third, as we discussed before, in principle we can solve for the stochastic sequential equilibrium under any realization of future temperature rise. We do so for a particular realization; the median temperature path for which we have solved the deterministic sequential equilibrium. This choice ensures that as uncertainty about the path converges to zero, the model with uncertainty will deliver the same allocation as the model with deterministic fundamentals, thereby isolating the role of uncertainty.[23]

We obtain the deviation in the productivity path from the median temperature scenario by evaluating the simulated temperature path described above using equation (12). To take into account parameter uncertainty, in each simulation, we draw $\alpha_j^1$ and $\alpha_j^2$ from their asymptotic distribution according to the standard errors of the estimates.

Figure 7 shows the variance across simulations of the deviations in log productivity from the deterministic path with climate change. As an illustration, the solid lines show the variance of the simulations from the year 2014, and the dotted lines show the variance of the simulations from the year 2060. Two observations emerge. First, the variance gradually increases, reflecting the fanning-out pattern in Panel (a) of Figure 4. Second, even though future temperature rises are similarly uncertain across locations according to Figure 4, the variance in future productivity is much higher in the South and in agriculture than in other regions and other sectors. This spatial and sectoral heterogeneity arises because the higher temperature in the South and the damage function in agriculture imply a larger marginal effect of temperature rise on productivity in the South and in agriculture than in other regions and sectors. Thus, when the uncertainty in temperature/parameters is introduced, the variance of the change in productivity increases more there.

Turning to the results, the dashed lines in Figure 5 show the evolution of the population in each broad region with uncertainty. Population moves away from southern locations, and especially from the Southwest, at a faster rate compared to the scenario with deterministic fundamentals. This finding reflects that the South faces larger uncertainty about the productivity path, which makes it less attractive to risk-averse individuals compared to a deterministic world. Overall, the Southwest experiences an additional 0.2 percentage points decline in population due to uncertainty, adding to the 3.8 percentage points decline in the scenario with deterministic fundamentals.

Considering uncertainty also affects welfare. Figure 8 plots the change in households' welfare due to the introduction of uncertainty to the deterministic climate change scenario. Since individuals are risk averse, their welfare decrease everywhere. The welfare decline tends to be more pronounced in southern and northeastern regions than in other regions. Aggregate welfare loss

[23] Although equation (13) implies current realization affects future distribution, under the law of iterative expectations, along the median temperature path, agents' expected future path will be the same as the future median temperature.

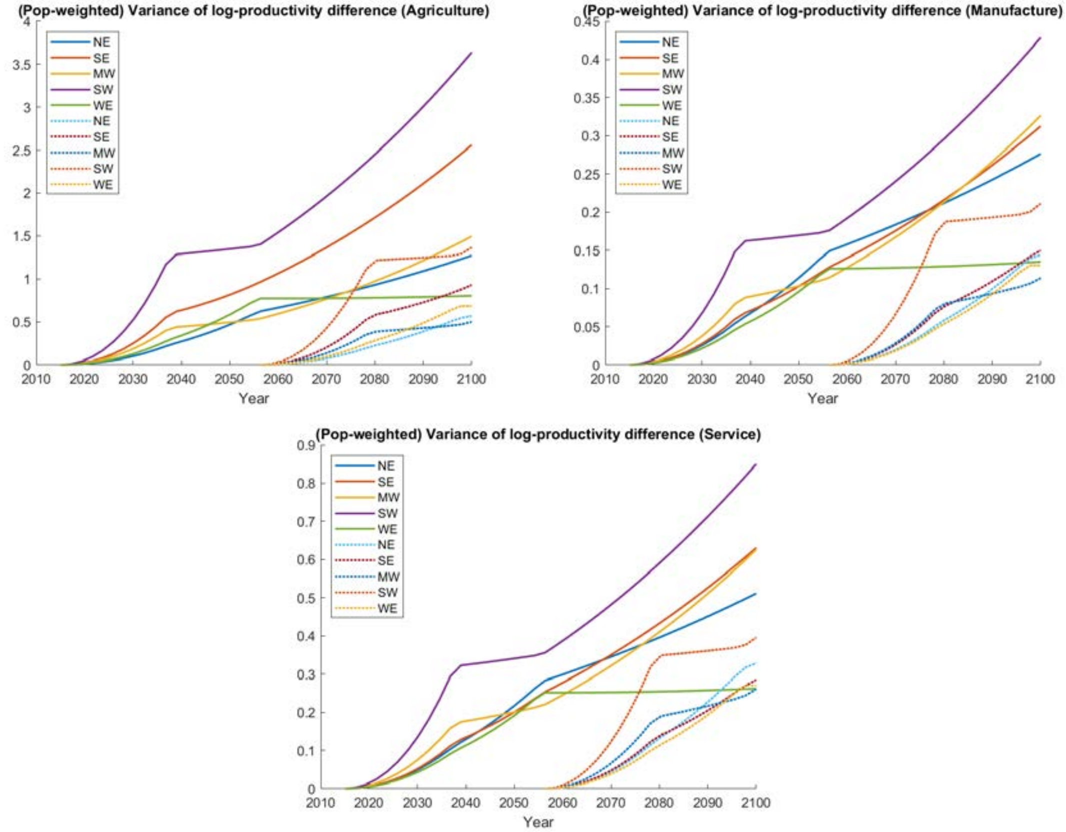



Figure 7: Uncertainty about Future Productivity

Notes: Each panel plots the variance of log productivity deviation in each sector (left: agriculture, middle: manufacture, right: service) used in simulations. The solid lines are for simulations from 2014; the dotted lines are for simulations from 2060.

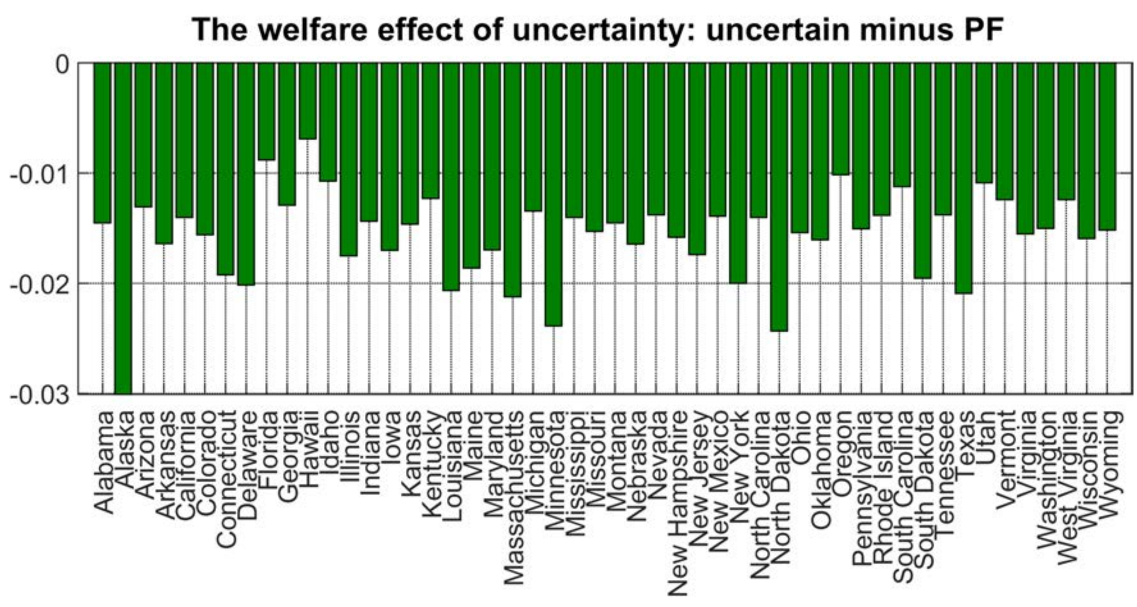



Figure 8: The Welfare Effect of Uncertainty

Notes: This figure plots the percentage difference in the welfare of the current residents in the presence of uncertainty compared to in the absence of it.

due to the uncertainty in climate change alone is about 1.1%. Adding this to the 0.9% welfare loss from perfectly anticipated climate change brings the welfare loss of climate change to 2%.

### 3.4 The Role of Disagreement

We now consider another departure from perfect foresight—heterogeneous beliefs among agents.

In Figure 1b we illustrated the spatial heterogeneity in the beliefs about climate change in 2021.[24] In this section, we incorporate such heterogeneous belief by assuming that in 2014, an exogenous share of the population in each state does not believe in climate change, and thinks the future temperature will be the same as that in 2014. The remaining people believe in climate change and hold the belief that future productivity will evolve according to the median temperature pace. We measure these shares using the 2014 vintage of the survey underlying Figure 1b. We also assume there is no uncertainty for either the believers or the skeptics.

We apply Proposition 4 to solve for the dynamic allocations with heterogeneous beliefs. In particular, we solve for the counterfactual path of this economy under Assumption 1, where each agent type believes that the other type will be convinced by the data in the next period, as deviations from the baseline scenario in which everyone is a believer.

Figure 9 plots the reallocation effect in the presence of climate skeptics in red dashed lines (the left axis). Compared to the deterministic baseline scenario in which everyone believes in climate change, the presence of climate skeptics results in a higher share of the population located in the South and a slower relocation of the population across space. Intuitively, as a fraction of the population does not believe the temperature will rise, they tend to stay in or move to the South. Consistent with this intuition, the yellow solid lines (the right-axis) shows that generally, the share of believers decreases in the Southeast and Southwest and increase in the Northeast.

Figure 10 plots the change in the welfare of believers in the presence of climate skeptics relative to the baseline scenario. We find that believers in the North and Midwest tend to be better off, whereas believers currently in the South tend to be worse off. The increase in the welfare of believers in Northern locations stems from the more muted migration to the North, which reduces the labor market competition faced by workers there; the same force leads to a decrease in the welfare of believers in the South. The change in welfare is noticeable for some states; for example, 1.6 percentage points in Maine and 0.5 percentage points in Wisconsin. At the aggregate, the welfare of believers increases by 0.08%.

In sum, in this section, we have applied our methodology to study the impact of climate change on welfare and population relocation across space within the United States under uncertainty and heterogeneous beliefs about the path of climate change. We find that introducing uncertainty alone leads to a 1.1 percent aggregate welfare loss in the United States, larger than the effect of perfectly foreseen climate change. The larger welfare losses are shaped by the risk aversion of individuals and by the significant impact of uncertainty in southern locations. Uncertainty also leads to faster spatial reallocation of the population to the North and Midwest. On the other hand,

[24] This disagreement might be correlated with other observables, such as age or skill, which our model abstracts from. Our methodology works in a similar way if such individual-level heterogeneity is considered.

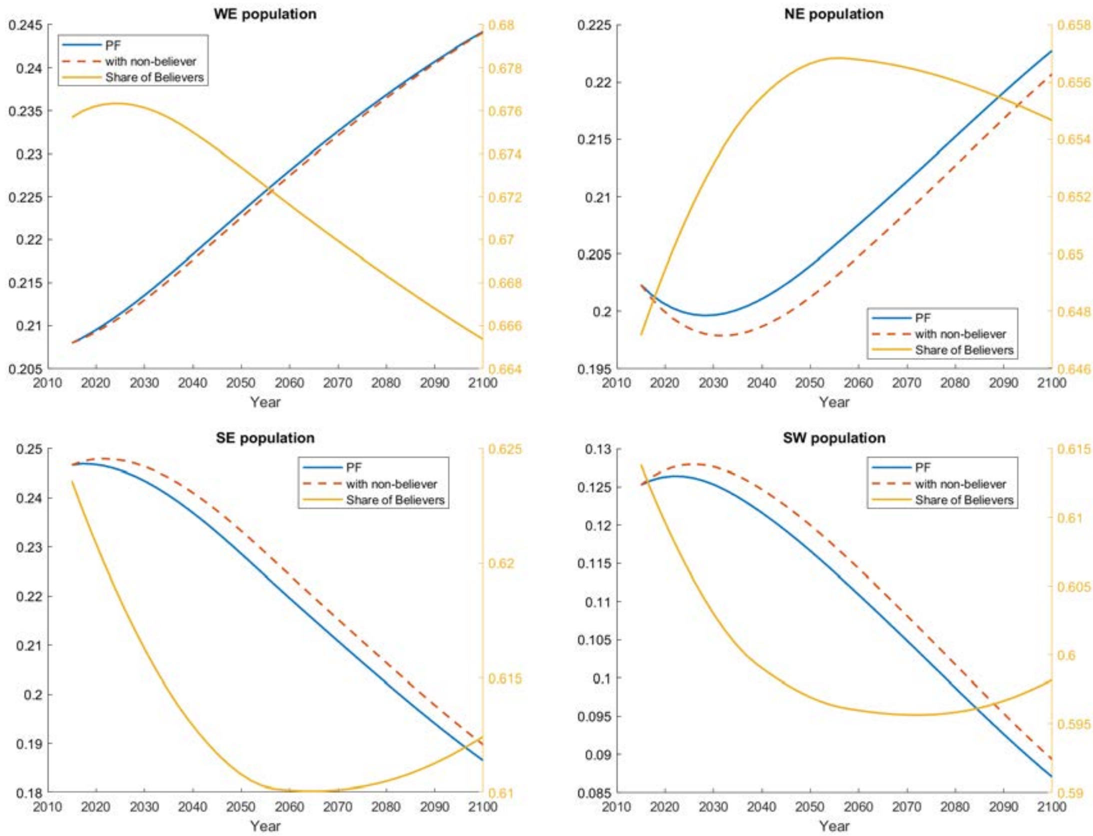



Figure 9: Reallocation in the Presence of Nonbelievers

Notes: The figure shows the share of each region's share of the population and share of believers over time under different scenarios.

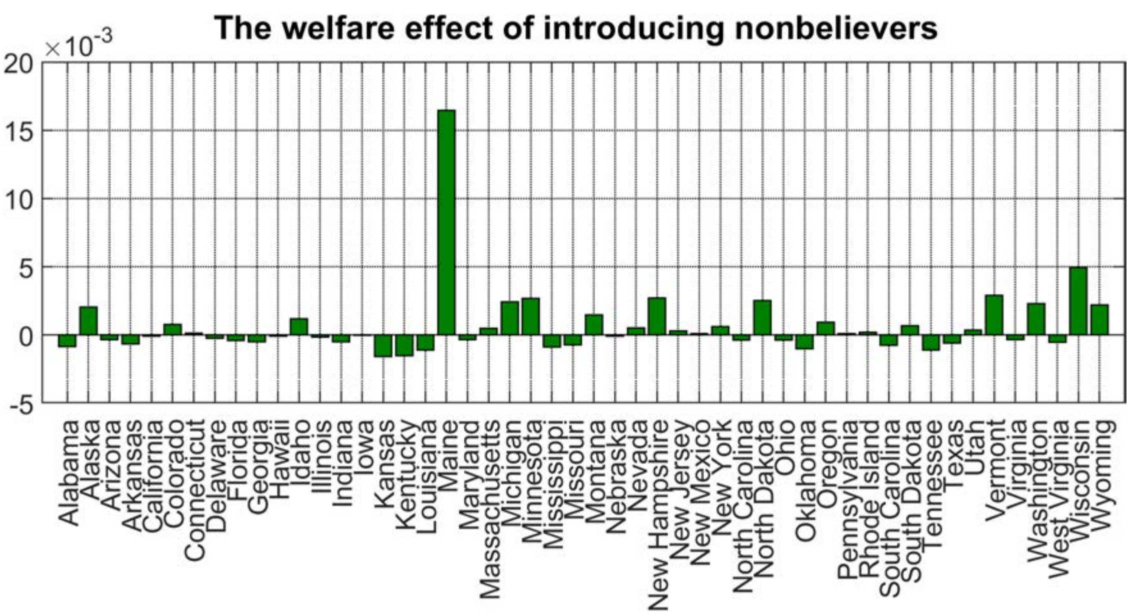



Figure 10: The Welfare of Believers Compared to PF

This figure plots the percentage difference in the welfare of believers in the presence of nonbelievers to the welfare in the baseline model where all agents have perfect foresight.

the presence of climate skeptics reduces the welfare of other people in the South, increases the welfare of other people in the North, and slows down spatial relocation.

## 4 Ex-post Application: the China Shock

Our second application studies ex-post the spatial effects of changes in fundamental productivity in the presence of evolving (and possibly biased) beliefs. We apply our methodology to study the impact of rapid productivity growth in China on the manufacturing employment share and welfare of workers in the United States.

As is well known, after China's accession to the WTO in 2001, its exports to the United States accelerated. The rapid growth in Chinese exports to the United States might have been surprising to many observers in the U.S., especially since China received the most-favored-nation treatment before 2001 (Handley and Limão, 2017). We model the source of import growth as driven by the catch-up in China's fundamental productivity with that of the United States. We allow workers in the U.S. to form beliefs ($f(z_{t+1}|z^t)$) about the catch-up process ($g(z_{t+1}|z^t)$) and will recover these beliefs by matching the employment projection at each moment in time. The belief processes and the actual allocation over the 'China shock' period will also allow us to construct the perfect foresight path as described in Section 2.5 and the counterfactual path without the China shock using Proposition 1.

### 4.1 Setup and Data

As in the climate change application, we extend the baseline model to incorporate multiple sectors and intermediate goods. Since imports from China are mostly manufacturing goods, we incorporate two sectors—manufacturing and others—in the model. Our analysis will include states in the United States and 28 other countries. We allow for migration across sectors and states within the United States, but not internationally. Following many existing quantitative studies of the China shock, we assume flow utility $U(c_{it}) = \log(c_{it})$.[25]

We assume that agents have perfect foresight about all parameters except for China's productivity, the premise being not many people in 2001 anticipated China to catch up as quickly as it did.[26] We will use the projections in Figure 1a to discipline these beliefs, implicitly assuming that if people had perfectly anticipated the productivity path of China, their projections about manufacturing employment shares would have been aligned with the actual shares. This choice is primarily motivated by the growing research that emphasizes the role of trade with China in the decline in U.S. manufacturing activity (Autor et al., 2013; Pierce and Schott, 2016; Handley and Limão, 2017), but it is also consistent with the measure productivity dynamics between the two countries, which shows the catch-up in productivity is mostly due to growth in China rather than abrupt changes in the United States.[27]

[25]We set $\beta = 0.96$ and use the same migration and trade elasticities as in the climate change application.

[26]The very use of the term 'China shock' in the trade literature is consistent with this premise.

[27]Alternatively, we can assume that people expected U.S. manufacturing productivity to grow faster than it did. We

To recover the perfect foresight and counterfactual paths from the actual allocation, we start by constructing the time series of the actual migration, trade, and production data. As most of these data sources have been described in Section 3.1, here we only describe additional steps that are necessary to construct the time series. First, the CFS covers only 2002, 2007, 2012, and 2017. We interpolate inter-state manufacturing trade between 2000 and 2016 using the information from the nearest covered year. Second, the job-to-job flow data that we use to construct the migration matrix misses some states as either the destination or the origin over the period 2000-2009, when the dataset gradually expanded coverage to all states. We fill in the missing values by scaling future observations of each cell backward using the change in migration over time for that cell from Caliendo et al. (2019). Third, to inform the belief process, we need to measure the productivity of China and other countries before the year 2000, for which we use international trade and price data. We obtain price data from the Penn World Table. Since WIOD does not extend long enough before 2000, we obtain trade flows from EORA.

### 4.2 Productivity Catch-up and Evolving Beliefs

We first use a model inversion to recover the time series of manufacturing productivity in China and the United States. In our model, domestic expenditure shares, wages, and price indices satisfy the following relationship:

$$\pi_{nn,j}^{t} = z_{nj}^{t} \left( w_{nj}^{t} / P_{nj}^{t} \right)^{-\theta \cdot \gamma_{nj}},$$

where $\gamma_{nj}$ are the value-added shares in production, $w_{nj}^{t}$ is wage, and $P_{nj}^{t}$ is the import price index from the Penn World Table. We can therefore measure $z_{nj}^{t}$ using the data at hand.

The left panel of Figure 11 plots the $\log(z_{CHN,M}^{t}) - \log(z_{USA,M}^{t})$ over time. We can see in the figure that in the first ten years of the sample, China experienced some catch-up in relative productivity and then gave away part of the progress at the end of the 1990s. However, starting from 2000-2001, the year of China's accession to the WTO, the catch-up process accelerated. By 2016 the log difference in fundamental productivity with the United States shrinks to approximately 0.5, which is a quarter of the log difference in the year 2000.

Motivated by this S-shaped pattern in relative productivity after 2000, we model the catch-up process as a logistic function given by

$$\log(z_{CN,M}^{t}) - \log(z_{US,M}^{t}) = \alpha_3 \cdot \frac{\exp\left(\alpha_1 * t\right)}{\exp\left(\alpha_1 * t\right) + \alpha_2}, \tag{14}$$

where the coefficients $\alpha_1$, $\alpha_2$, and $\alpha_3$, govern the catch-up speed, the initial productivity gap, and the long-run productivity of China relative to the United States, respectively. We estimate this specification using the data over the period 2000-2016. We estimate that $\alpha_1 = -0.22$, $\alpha_2 = 0.11$, and $\alpha_3 = -2.4$. The solid line in the left panel of Figure 11 shows the fit. While the model implies a smoother catch-up process than the data, overall it fits the data well.

view this alternative as less appealing given the lack of trend breaks in the U.S. productivity series.

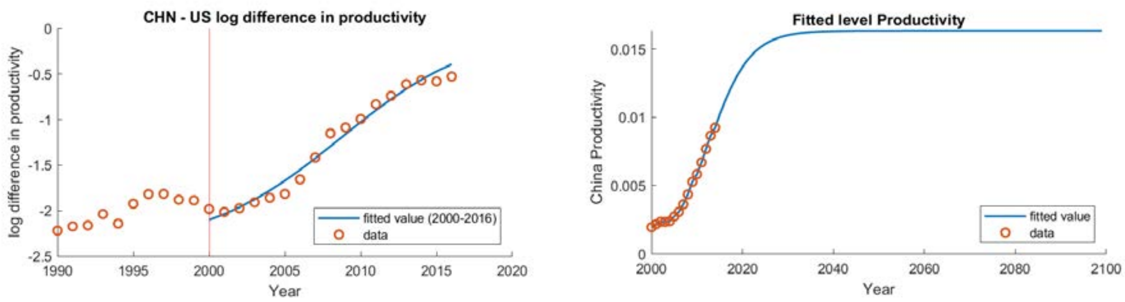



Figure 11: Measured Manufacturing Productivity: U.S. and China
The left panel plots the log difference between manufacturing productivity in China and the U.S. from 1990 to 2016 and the fitted line based on data from 2000 to 2016. The right panel plots the level of China's manufacturing productivity according to the fitted model in and out of the sample period.

We treat the estimated empirical catch-up model as the true process governing the manufacturing productivity of China relative to that of the United States. Assuming that the U.S. productivity is constant after 2016 and applying the estimated process, we obtain the level of manufacturing productivity in China, which is shown in the right panel of Figure 11. Under this specification, China's productivity grows until 2040, after which it reaches the steady state value determined by $\alpha_3$. This solid line is essentially the $g(z_{t+1}|z^t)$ in the theoretical section.

Forward-looking agents form beliefs about the catch-up speed, captured by $\alpha_1$, and update their beliefs as new data arrive. We recover agents' belief about $\alpha_1$—and by implication, their belief about the fundamental productivity in China, $f(z_{t+1}|z^t)$—for the period 2001-2006 by implementing the following algorithm:

1. Choose agents' beliefs about $\alpha_1$ in each of the following years: 2002, 2004, and 2006, interpolating the beliefs for the years in between.[28]

2. Given beliefs about $\alpha_1$, generate beliefs about China's productivity using the estimated catch-up process.

3. Use the algorithm described in Section 2.5 backward from 2007 to 2001, along the way recovering agents' beliefs about manufacturing employment share in each year.

4. Check if the model-implied manufacturing employment shares projection matches the empirical projection; if not, go back to step 1 and revise the belief parameters.

[28] We choose the belief parameter for only these even-number years because employment projection from the BLS is only available for these years. We assume that starting from 2008, agents in the economy have perfect foresight. This assumption is motivated by the observation that in Figure 1a, by 2006-2008, which reflects the end of the China shock period, the employment projection about the future is already largely aligned with the data. We note that our methodology does not need to impose that agents are perfect foresight when the data ends. We can always first solve for a sequential path according to agents' belief from the end of the data, from which it is straightforward to recover the perfect foresight path, as we discussed in detail in the theory section.

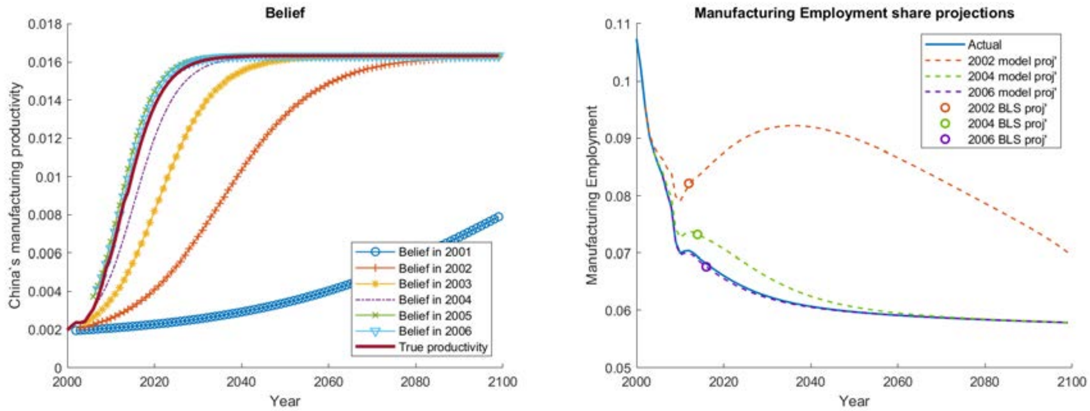



Figure 12: Employment Projection and Agent's Belief

Note: The left panel presents the recovered beliefs about China's manufacturing productivity level formed at a particular year (solid lines differentiated by color) and the actual productivity process (blue dashed line). The right panel presents the actual manufacturing employment share (blue solid line), the model implied manufacturing employment share projection starting from 2002, 2004, and 2006 (dashed lines), and the BLS 10-year projections (circles) from the corresponding years.

Figure 12 illustrates the outcomes from this algorithm. The left panel presents the recovered beliefs about the level of China's manufacturing productivity. Each color represents the belief formed in a particular year. The brown solid line is the actual process for productivity. In the initial years, agents are too pessimistic about China's productivity. Gradually, they revise their belief, and by 2006, the belief is almost fully aligned with actual productivity. The right panel shows that under this belief, the model implied manufacturing employment share projection (dashed lines) is consistent with the BLS 10-year projection (represented using circles).

### 4.3 The Impact of the China Shock under Evolving Beliefs

With the belief process described in the previous subsection, we apply the methodology described in Section 2.5 to study the impact of the China shock. We think of the 'shock' as the rapid productivity catch-up of China that might have been surprising to many observers. Accordingly, we define the counterfactual no-shock scenario as one in which China's manufacturing productivity grows at a slower rate (e.g., 3.5% per annum) until reaching its steady-state value, and in which agents perfectly foresee this slower catch-up. We answer two questions: first, how would the U.S. economy have evolved with a slower (and perfectly anticipated) China productivity catch-up? The difference between this counterfactual path and the actual data reflects both the faster productivity growth in China and the estimated beliefs about the evolution of China's productivity. Since the bulk of the literature using dynamic spatial models has assumed that the data are generated by agents with perfect foresight, our second question is: how different will our answer be, if we have imposed that in the data agents have perfect foresight?

Figure 13 compares the counterfactual productivity processes under different assumptions on the growth rates with the actual one. We use a 3.5% per annum growth rate, which implies that productivity in China reaches the steady state in 2060, about 30 years later than implied by the true catch-up process.

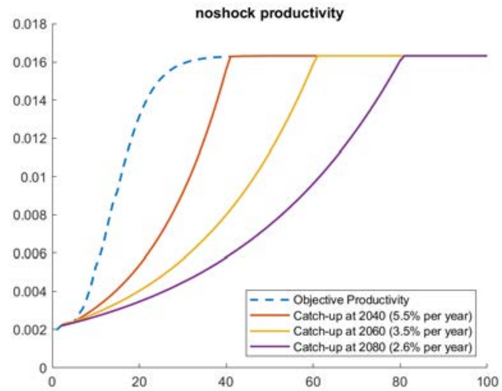



Figure 13: Counterfactual China Productivity Growth

Note: This figure presents the true productivity process (dashed blue line) with three counterfactual productivity processes for China's manufacturing productivity. The first counterfactual productivity process is that China's manufacturing productivity grows at 5.5% per annum, reaching the peak in 2040 (red solid line), second counterfactual productivity process is that it grows at 3.5% per annum, reaching the peak in 2060 (yellow solid line), and the third counterfactual productivity process is that it grows at 2.6% per annum, reaching the peak in 2080 (purple solid line).

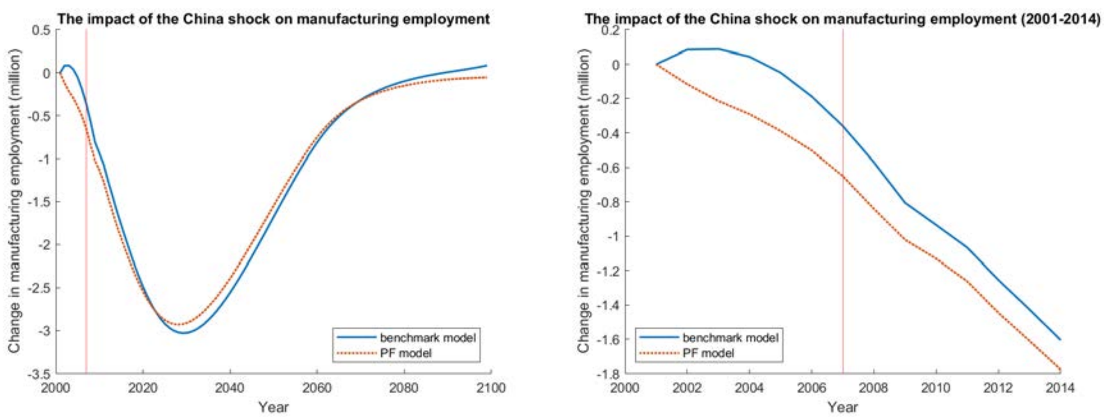



Figure 14: The Employment Effect of the China Shock

Note: The left panel presents the effects of the China shock measured as the change in manufacturing employment between the economy with China shock and the economy without China shock (China's manufacturing productivity grows at 3.5% per annum) under the perfect foresight model (PF model) and our benchmark model. The right panel focuses on the data period: from 2001 to 2014.

The red solid line in Figure 14 presents the impact of the China shock, defined above, on U.S. manufacturing employment. We can see that total manufacturing employment decreases over the first few years, reaching a peak decrease of approximately three million, before it gradually converges back. This reversion in the path of manufacturing employment results from the fact that in the long run, both in the baseline economy and in the counterfactual economy, China's manufacturing productivity eventually reaches the same level.

Regarding our second counterfactual question, the red dashed line in Figure 14 displays the inferred effect on manufacturing employment if we had assumed that the data are generated by agents with perfect foresight. On the left panel, we can see how the path of manufacturing employment under perfect foresight converges to the baseline economy as China's productivity catches up with the United States, in line with our discussion in the previous paragraph. However,

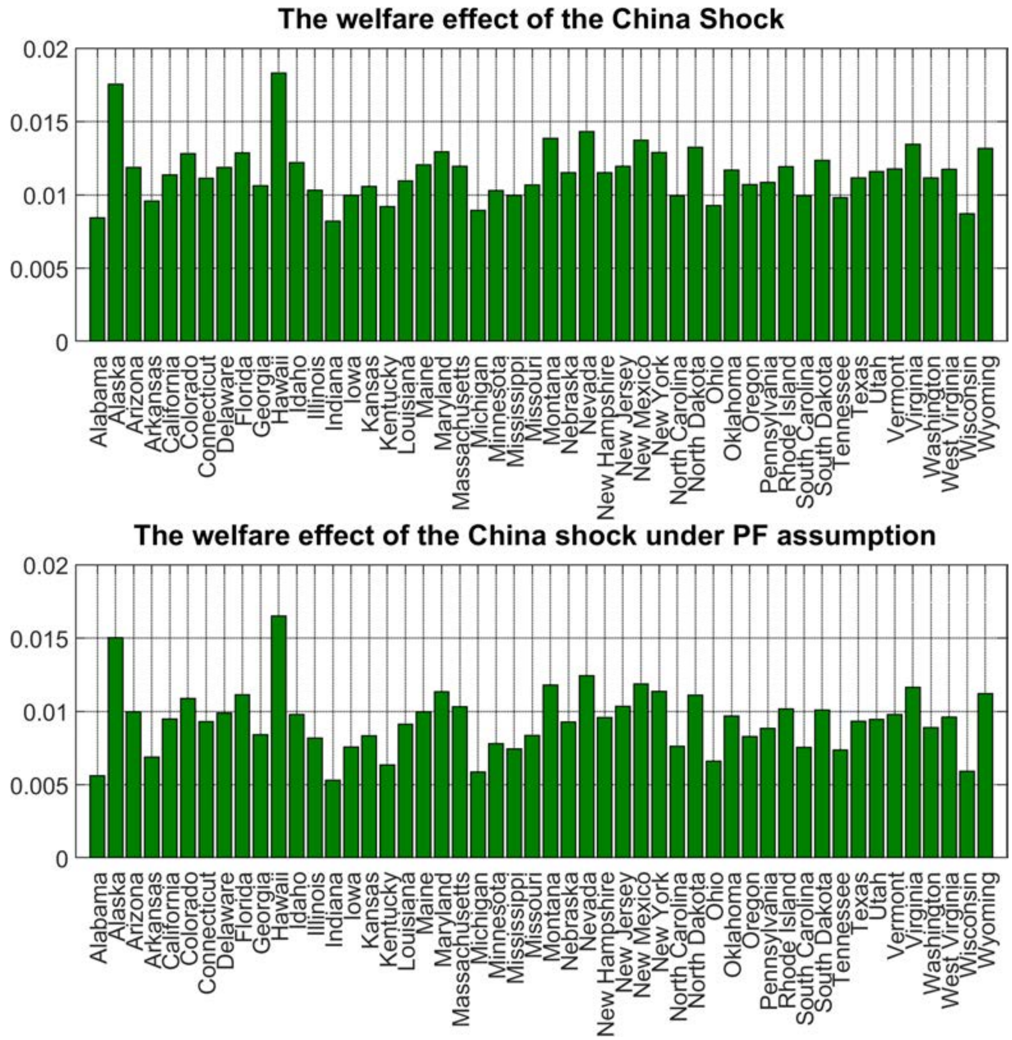



Figure 15: The Welfare Effect of the China Shock (percent)

Note: The upper figure presents the percentage change in welfare in each state under our benchmark model, and the bottom figure presents the percentage change in welfare in each state under the perfect foresight model.

zooming into the catch-up process, shown on the right Panel of Figure 14, we can see our model and the perfect foresight model find different effects. In particular, the perfect foresight model infers immediate reallocation out of manufacturing as workers perfectly anticipate productivity growth. On the other hand, in our benchmark model with evolving beliefs, agents initially did not anticipate productivity growth in China, so they did not reallocate from manufacturing. Subsequently, as agents update their beliefs (recall that agents' beliefs became aligned with the actual evolution by 2006-2008), manufacturing employment declines and the gap with the manufacturing decline under perfect foresight shrinks. However, even by 2014, the difference in the decline of manufacturing employment between our model and the perfect foresight model remains more than 10%.

Figure 15 reports the welfare effect of the China shock on U.S. states in our model and the perfect foresight model. On average, Our model implies around 1.12% gains in welfare, around one-tenth higher than under perfect foresight (1.04%). The difference in the inferred welfare gains is heterogeneous across space. For example, in Alabama, our model infers 0.8% welfare gains, about one-half higher than the perfect foresight model. One might have the intuition that perfect foresight enables agents to make more informed decisions, thereby improving their welfare, but it is important to note that this exercise is not about the value of information. Instead, it is about how researchers making different assumptions will infer different welfare effects. In our stochastic model with evolving beliefs, observed decisions are made under overly pessimistic beliefs about China's productivity, so we implicitly rationalize the actual migration rates with lower mobility costs than in the perfect foresight environment, which in turn leads to higher welfare gains from the China shock. In addition, evolving beliefs also imply a slower transition out of manufacturing than the perfect foresight model, as we discussed before. The decrease in manufacturing imports under this slower manufacturing employment reallocation leads to improvements in the U.S. terms of trade, further increasing welfare.

## 5 Conclusion

Agent's anticipation about the future path of fundamentals plays an important role in shaping the impact of shocks in dynamic spatial environments. In this paper, we have built a generic dynamic spatial equilibrium model with stochastic fundamentals. We develop a local approximation method to solve this model for counterfactual analysis, allowing agents' beliefs about these fundamentals to be evolving, uncertain, and heterogeneous among groups of agents. We further extend our method to conduct counterfactuals and disentangle the role of beliefs and the changes in model fundamentals in shaping economic outcomes, in scenarios where the researchers observe a sequence of data generated by an economy with imperfect beliefs and other time-varying fundamentals but do not observe the fundamentals directly.

To showcase our methodology, we apply it to two settings: an ex-ante study of the spatial reallocation in response to climate change; an ex-post study on the impact of China's manufacturing productivity growth on U.S. employment. In both cases, evidence demonstrates the departures of agents' beliefs from perfect foresight, and our quantitative applications demonstrate the importance of incorporating such departures.

## References

Artuç, Erhan, Shubham Chaudhuri, and John McLaren, "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 2010, *100* (3), 1008–45.

Autor, David H, David Dorn, and Gordon H Hanson, "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–2168.

Balboni, Clare, "In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities," 2019.

Bilal, Adrien, "Solving heterogeneous agent models with the master equation," 2022.

Bombardini, Matilde, Bingjing Li, and Francesco Trebbi, "Did US Politicians Expect the China Shock?," *American Economic Review*, 2023, *113* (1), 174–209.

Bui, Ha, Zhen Huo, Andrei A Levchenko, and Nitya Pandalai-Nayar, "Information Frictions and News Media in Global Value Chains," 2022.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 2015, *82* (1), 1–44.

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

Cogley, Timothy and Thomas J Sargent, "The conquest of US inflation: Learning and robustness to model uncertainty," *Review of Economic dynamics*, 2005, *8* (2), 528–563.

Cruz, José-Luis, "Global warming and labor market reallocation," Technical Report, Working Paper 2023.

_ and Esteban Rossi-Hansberg, "The economic geography of global warming," *Review of Economic Studies, Forthcoming*, 2023.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Unbalanced trade," *American Economic Review*, 2007, *97* (2), 351–355.

Desmet, Klaus, Robert E Kopp, Scott A Kulp, Dávid Krisztián Nagy, Michael Oppenheimer, Esteban Rossi-Hansberg, and Benjamin H Strauss, "Evaluating the Economic Cost of Coastal Flooding," *American Economic Journal: Macroeconomics*, 2021, *13* (2), 444–486.

Dickstein, Michael J and Eduardo Morales, "What do exporters know?," *The Quarterly Journal of Economics*, 2018, *133* (4), 1753–1801.

Dix-Carneiro, Rafael, João Paulo Pessoa, Ricardo M Reyes-Heroles, and Sharon Traiberman, "Globalization, trade imbalances and labor market adjustment," 2021.

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

Fujiwara, Thomas, Eduardo Morales, and Charly Porcher, "A revealed-preference approach to measuring information frictions in migration decisions," 2020.

Handley, Kyle and Nuno Limão, "Policy uncertainty, trade, and welfare: Theory and evidence for China and the United States," *American Economic Review*, 2017, *107* (9), 2731–2783.

Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A Sims, "Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models," *Journal of Economic Dynamics and Control*, 2008, *32* (11), 3397–3414.

Kleinman, Benny, Ernest Liu, and Stephen J Redding, "Dynamic spatial general equilibrium," *Econometrica*, 2023, *91* (2), 385–424.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran, "Scarring body and mind: the long-term belief-scarring effects of Covid-19," 2020.

Kreps, David M, "Anticipated utility and dynamic choice," *Econometric Society Monographs*, 1998, *29*, 242–274.

Nordhaus, William, *Managing the global commons: the economics of climate change*, The MIT Press, 1994.

Pang, Xinle and Sun Pin, “Moving into Risky Floodplains: The Spatial Implications of Federal Disaster Relief Policies,” *Unpublished working paper*, 2022.

Pierce, Justin R and Peter K Schott, “The surprisingly swift decline of US manufacturing employment,” *American Economic Review*, 2016, *106* (7), 1632–1662.

Rodríguez-Clare, Andrés, Mauricio Ulate, and Jose P Vasquez, “Trade with nominal rigidities: Understanding the unemployment and welfare effects of the china shock,” 2020.

Rudik, Ivan, Gary Lyn, Weiliang Tan, and Ariel Ortiz-Bobea, “The Economic Effects of Climate Change in Dynamic Spatial Equilibrium,” 2022.

Schmitt-Grohé, Stephanie and Martın Uribe, “Solving dynamic general equilibrium models using a second-order approximation to the policy function,” *Journal of Economic Dynamics and Control*, 2004, *28* (4), 755–775.

Townsend, Robert M, “Forecasting the forecasts of others,” *Journal of Political Economy*, 1983, *91* (4), 546–588.

# Online Appendix: Learning and Expectations in Dynamic Spatial Economies

Jingting Fan
Penn State

Sungwan Hong
Penn State

Fernando Parro
Penn State and NBER

# Appendix A Theory

This appendix presents the proofs of the propositions presented in the main text and additional derivations. Throughout, we define $\hat{v}_{it+1}(z^t) \equiv v_{it+1}(z^t) - \bar{v}_{it+1}$ and $\hat{v}_{it}(z^t) \equiv v_{it}(z^t) - \bar{v}_{it}$. For all other exogenous or endogenous variables $x$, we define $\hat{x}(z^t) \equiv \log \hat{x}(z^t) - \log \bar{x}$. To simplify the notation, we denote a conditional expectation by $\mathbb{E}_t x_{nt'} \equiv \mathbb{E}(x_{nt'}|z^t)$. We assume the process of trade cost $\kappa$ is perfectly foresighted. However, we keep the terms related to trade cost in each equation to show the proofs go through when other fundamentals are stochastic.

## A.1 Proof of Proposition 1

We prove the proposition by deriving the first-order approximation of equations (2)-(7). Algebras are relegated to Appendix A.6.1.

$$
\begin{aligned}
\hat{v}_{nt}(z^t) &= \frac{\partial U(\bar{\varpi}_{nt}, \bar{P}_{nt})}{\partial \log \bar{\varpi}_{nt}} \hat{\varpi}_{nt}(z^t) + \frac{\partial U(\bar{\varpi}_{nt}, \bar{P}_{nt})}{\partial \log \bar{P}_{nt}} \hat{P}_{nt}(z^t) + \beta \sum_{i=1}^{N} \bar{\mu}_{nit} \mathbb{E}_t \hat{v}_{it+1} && \text{(A.1)}\\
\hat{\mu}_{nit}(z^t) &= \frac{\beta}{\nu} \sum_{k=1}^{N} \Big( \mathbb{1}(k=i) - \bar{\mu}_{nkt} \Big) \mathbb{E}_t \hat{v}_{kt+1} && \text{(A.2)}\\
\hat{l}_{nt+1}(z^t) &= \sum_{i}^{N} \frac{\bar{\mu}_{int} \bar{l}_{it}}{\bar{l}_{nt+1}} \Big( \hat{\mu}_{int}(z^t) + \hat{l}_{it}(z^{t-1}) \Big) && \text{(A.3)}\\
\hat{\lambda}_{nit}(z^t) &= -\theta(\hat{w}_{it}(z^t) + \hat{\kappa}_{int} - \hat{P}_{nt}(z^t)) + \hat{z}_{it} && \text{(A.4)}\\
\hat{P}_{nt}(z^t) &= \sum_i \bar{\lambda}_{nit} \Big( \hat{w}_{it}(z^t) + \hat{\kappa}_{int} - \frac{1}{\theta} \hat{z}_{it} \Big) && \text{(A.5)}\\
\hat{w}_{nt}(z^t) + \hat{l}_{nt}(z^{t-1}) &= \sum_i \frac{\bar{\lambda}_{int} \bar{\varpi}_{it} \bar{l}_{it}}{\bar{\varpi}_{nt} \bar{l}_{nt}} \Big( \hat{\lambda}_{int}(z^t) + \hat{w}_{it}(z^t) + \hat{l}_{it}(z^{t-1}) \Big) && \text{(A.6)}\\
\text{For } t' > t, & && \text{(A.7)}\\
\mathbb{E}_t \hat{v}_{nt'} &= \frac{\partial U(\bar{\varpi}_{nt'}, \bar{P}_{nt'})}{\partial \log \bar{\varpi}_{nt'}} \mathbb{E}_t \hat{\varpi}_{nt'} + \frac{\partial U(\bar{\varpi}_{nt'}, \bar{P}_{nt'})}{\partial \log \bar{P}_{nt'}} \mathbb{E}_t \hat{P}_{nt'} + \beta \sum_{i=1}^{N} \bar{\mu}_{nit'} \mathbb{E}_t \hat{v}_{it'+1} && \text{(A.8)}\\
\mathbb{E}_t \hat{\mu}_{nit'} &= \frac{\beta}{\nu} \sum_{k=1}^{N} \Big( \mathbb{1}(k=i) - \bar{\mu}_{nkt'} \Big) \mathbb{E}_t \hat{v}_{kt'+1} && \text{(A.9)}\\
\mathbb{E}_t \hat{l}_{nt'+1} &= \sum_{i}^{N} \frac{\bar{\mu}_{int'} \bar{l}_{it'}}{\bar{l}_{nt'+1}} \Big( \mathbb{E}_t \hat{\mu}_{int'} + \mathbb{E}_t \hat{l}_{it'} \Big) && \text{(A.10)}\\
\mathbb{E}_t \hat{\lambda}_{nit'} &= -\theta(\mathbb{E}_t \hat{w}_{it'} + \mathbb{E}_t \hat{\kappa}_{int'} - \mathbb{E}_t \hat{P}_{nt'}) + \mathbb{E}_t \hat{z}_{it'} && \text{(A.11)}\\
\mathbb{E}_t \hat{P}_{nt'} &= \sum_i \bar{\lambda}_{nit'} \Big( \mathbb{E}_t \hat{w}_{it'} + \mathbb{E}_t \hat{\kappa}_{int'} - \frac{1}{\theta} \mathbb{E}_t \hat{z}_{it'} \Big) && \text{(A.12)}\\
\mathbb{E}_t \hat{w}_{nt'} + \mathbb{E}_t \hat{l}_{nt'} &= \sum_i \frac{\bar{\lambda}_{int'} \bar{\varpi}_{it'} \bar{l}_{it'}}{\bar{\varpi}_{nt'} \bar{l}_{nt'}} \Big( \hat{\lambda}_{int'} + \hat{w}_{it'} + \hat{l}_{it'} \Big). && \text{(A.13)}
\end{aligned}
$$

This system of linear equations requires (i) the approximation points $\{\bar{\mu}_{nit'}, \bar{l}_{nt'}, \bar{\lambda}_{nit'}, \bar{\varpi}_{nt'}, \bar{P}_{nt'}\}_{i=1,n=1,t'=t}^{N,N,T}$ from the perfect foresight path, (ii) the beliefs of agents in period $t$ on the expected values of future fundamentals as deviations from the fundamentals underlying the perfect foresight path $\mathbb{E}_t \hat{z}_{t'}$, and

(iii) the labor allocation in each region in period $t$, $\hat{l}_{it}$.

Note that the labor allocation in period $t$ is determined in period $t-1$ and therefore the deviation of expected values of future fundamentals from the perfect foresight path in period $t$ does not affect the labor allocation at $t$, i.e., $\{\mathbb{E}_t \hat{l}_{nt}\}_{n=1}^{N} = 0$. Given this, we can solve for $\left\{\hat{v}_{it}, \hat{\mu}_{nit}, \hat{l}_{it}, \hat{w}_{it}, \hat{\lambda}_{nit}, \hat{P}_{it}\right\}_{i=1,n=1}^{N,N}$, and we obtain $\{\mathbb{E}_t \hat{l}_{nt+1}\}_{n=1}^{N}$ from equation A.10. Given $\{\mathbb{E}_t \hat{l}_{nt+1}\}_{n=1}^{N}$, we can solve for $\{\mathbb{E}_t \hat{v}_{it+1}, \mathbb{E}_t \hat{\mu}_{nit+1}, \mathbb{E}_t \hat{w}_{it+1}, \mathbb{E}_t \hat{P}_{it+1}\}_{i=1,n=1}^{N,N}$. Iterate until period $T$ and we obtain $\{\mathbb{E}_t \hat{v}_{it'}, \mathbb{E}_t \hat{\mu}_{nit'}, \mathbb{E}_t \hat{w}_{it'}, \mathbb{E}_t \hat{P}_{it'}\}_{i=1,n=1,t'=t+1}^{N,N}$

## A.2 Proof of Proposition 2

We explain in detail the algorithm that was laid out in Section 2.5.

**Recovering the Perfect Foresight from the Observed Allocation.** Let the actual outcomes in period $t$ be denoted using a tilde, i.e., $\tilde{v}_{nt}$, $\tilde{w}_{nt}$, $\tilde{P}_{nt}$, $\tilde{\lambda}_{nit}$, $\tilde{\mu}_{nit}$, $\tilde{l}_{nt}$. Approximation points are denoted by bar and differ by time superscript for each system of equations explained below.

i Starting from period $T$, solve for the expected outcomes of the trade equilibrium in period $T$ according to the beliefs of agents in period $T-1$, namely $\mathbb{E}_{T-1} w_{iT}$, $\mathbb{E}_{T-1} P_{iT}$ and $\mathbb{E}_{T-1} \lambda_{inT}$ for all $i, n$ by approximations around the actual outcomes in period $T$. The input for this step is the actual outcome in $T$ and the deviations in agents' belief in period $T-1$ about the productivity in $T$ from the actual productivity, namely $\mathbb{E}_{T-1} \hat{z}_T$.

   Define $\mathbb{E}_{T-1} \hat{x}_T \equiv \mathbb{E}_{T-1} \log(x_T) - \log(\tilde{x}_T)$, for $x \in \{w, P, \lambda, l\}$. We use level difference for $v$, that is, $\mathbb{E}_{T-1} \hat{v}_T \equiv \mathbb{E}_{T-1} v_T - \tilde{v}_T$. Note that $\mathbb{E}_{T-1} l_{nT} = \tilde{l}_{nT}$ thus $\mathbb{E}_{T-1} \hat{l}_{nT} = 0$. Then, $\mathbb{E}_{T-1} w_T$, $\mathbb{E}_{T-1} P_T$, and $\mathbb{E}_{T-1} \lambda_T$ are obtained from the following system of equations:

$$\mathbb{E}_{T-1} \hat{\lambda}_{niT} = -\theta\left(\mathbb{E}_{T-1} \hat{w}_{iT} + \mathbb{E}_{T-1} \hat{\kappa}_{inT} - \mathbb{E}_{T-1} \hat{P}_{nT}\right) + \mathbb{E}_{T-1} \hat{z}_T \quad \text{(A.14)}$$

$$\mathbb{E}_{T-1} \hat{P}_{nT} = \sum_i \bar{\lambda}^T_{niT} \left(\mathbb{E}_{T-1} \hat{w}_{iT} + \mathbb{E}_{T-1} \hat{\kappa}_{inT} - \frac{1}{\theta} \mathbb{E}_{T-1} \hat{z}_T\right) \quad \text{(A.15)}$$

$$\mathbb{E}_{T-1} \hat{w}_{nT} + \mathbb{E}_{T-1} \hat{l}_{nT} = \sum_i \frac{\bar{\lambda}^T_{inT} \bar{w}^T_{iT} \bar{l}^T_{iT}}{\bar{w}^T_{nT} \bar{l}^T_{nT}} \left(\mathbb{E}_{T-1} \hat{\lambda}_{inT} + \mathbb{E}_{T-1} \hat{w}_{iT} + \mathbb{E}_{T-1} \hat{l}_{iT}\right) \quad \text{(A.16)}$$

   where the approximation points are coming from the actual outcomes, $(\bar{\lambda}^T_{inT}, \bar{w}^T_{nT}, \bar{P}^T_{nT}, \bar{l}^T_{nT}) = (\tilde{\lambda}_{inT}, \tilde{w}_{nT}, \tilde{P}_{nT}, \tilde{l}_{nT})$. Note that the level outcomes can be recovered by using the relationship $\mathbb{E}_{T-1} \log(x_T) = \mathbb{E}_{T-1} \hat{x}_T + \log(\tilde{x}_T)$ for $x \in \{w, P, \lambda, l\}$.

ii Append the output from the first step by $\tilde{\mu}_{niT-1}$, namely agents' *actual* migration decision in period $T-1$. As $\tilde{\mu}_{niT-1}$ is decided according to agents' belief in period $T-1$, together with the output from (i), it constitutes the solution to the agents' problem in period $T-1$, as defined in Proposition 1.
   From step 1 and 2, we obtain $\{\mathbb{E}_{T-1} \lambda_{niT}, \mathbb{E}_{T-1} P_{nT}, \mathbb{E}_{T-1} w_{nT}, \mathbb{E}_{T-1} l_{nT}, \mathbb{E}_{T-1} \mu_{niT-1}\}_{i=1,n=1}^{N,N}$

iii Solve for the expected outcomes in periods $\{T-1, T\}$ according to the beliefs of agents in period $T-2$ by approximating around the solution to agents' problem in period $T-1$, obtained from (ii). These outcomes include $\{\mathbb{E}_{T-2}\lambda_{nit}, \mathbb{E}_{T-2}P_{nt}, \mathbb{E}_{T-2}w_{nt}, \mathbb{E}_{T-2}l_{nt}, \mathbb{E}_{T-2}\mu_{nit}\}_{i=1,n=1,t=T-1}^{N,N,T}$. The input to this approximation is the output from (i) and (ii) and the deviations in agents' belief in period $T-2$ about the productivity in $\{T-1, T\}$ from their belief in period $T-1$ about it, i.e., $\mathbb{E}_{T-2}\log(z_t) - \mathbb{E}_{T-1}\log(z_t)$ for $t \in \{T-1, T\}$.

The outcomes can be obtained by solving the following system of equations for $t \in \{T-1, T\}$:

$$
\begin{aligned}
\mathbb{E}_{T-2}\hat{v}_{nt} &= \frac{\partial U(\bar{w}_{nt}, \bar{P}_{nt})}{\partial \log(\bar{w}_{nt})}\mathbb{E}_{T-2}\hat{w}_{nt} + \frac{\partial U(\bar{w}_{nt}, \bar{P}_{nt})}{\partial \log(\bar{P}_{nt})}\mathbb{E}_{T-2}\hat{P}_{nt} + \beta \sum_{i=1}^{N} \bar{\mu}_{nit}^{T-1}\mathbb{E}_{T-2}\hat{v}_{it+1} \\
\mathbb{E}_{T-2}\hat{\mu}_{nit} &= \frac{\beta}{\nu}\sum_{k=1}^{N}\left(\mathbb{1}(k=i) - \bar{\mu}_{nkt}^{T-1}\right)\mathbb{E}_{T-2}\hat{v}_{kt+1} \\
\mathbb{E}_{T-2}\hat{l}_{nt+1} &= \sum_{i}^{N}\frac{\bar{\mu}_{int}^{T-1}\bar{l}_{it}^{T-1}}{\bar{l}_{nt+1}^{T-1}}\left(\mathbb{E}_{T-2}\hat{\mu}_{int} + \mathbb{E}_{T-2}\hat{l}_{it}\right) \\
\mathbb{E}_{T-2}\hat{\lambda}_{nit} &= -\theta\left(\mathbb{E}_{T-2}\hat{w}_{it} + \mathbb{E}_{T-2}\hat{\kappa}_{int} - \mathbb{E}_{T-2}\hat{P}_{nt}\right) + \mathbb{E}_{T-2}\hat{z}_t \\
\mathbb{E}_{T-2}\hat{P}_{nt} &= \sum_{i}\bar{\lambda}_{nit}^{T-1}\left(\mathbb{E}\hat{w}_{it}(z^{T-2}) + \mathbb{E}_{T-2}\hat{\kappa}_{int} - \frac{1}{\theta}\mathbb{E}_{T-2}\hat{z}_t\right) \\
\mathbb{E}_{T-2}\hat{w}_{nt} + \mathbb{E}_{T-2}\hat{l}_{nt} &= \sum_{i}\frac{\bar{\lambda}_{int}^{T-1}\bar{w}_{it}^{T-1}\bar{l}_{it}^{T-1}}{\bar{w}_{nt}^{T-1}\bar{l}_{nt}^{T-1}}\left(\mathbb{E}_{T-2}\hat{\lambda}_{int} + \mathbb{E}_{T-2}\hat{w}_{it} + \mathbb{E}_{T-2}\hat{l}_{it}\right),
\end{aligned}
\tag{A.17}
$$

where the approximation points for $t = T-1$ period equations are coming from actual outcomes, i.e.,$(\bar{\lambda}_{niT-1}^{T-1}, \bar{P}_{nT-1}^{T-1}, \bar{w}_{nT-1}^{T-1}, \bar{l}_{nT-1}^{T-1}, \bar{\mu}_{niT-1}^{T-1}) = (\tilde{\lambda}_{niT-1}, \tilde{P}_{nT-1}, \tilde{w}_{nT-1}, \tilde{l}_{nT-1}, \tilde{\mu}_{niT-1})$, and the approximation points for $t = T$ period equations are coming from the outcomes we obtained from (i) and (ii), i.e., $(\bar{\lambda}_{niT-1}^{T-1}, \bar{P}_{nT-1}^{T-1}, \bar{w}_{nT-1}^{T-1}, \bar{l}_{nT-1}^{T-1}, \bar{\mu}_{niT-1}^{T-1})$ $= (\mathbb{E}_{T-1}\lambda_{niT}, \mathbb{E}_{T-1}P_{nT}, \mathbb{E}_{T-1}w_{nT}, \mathbb{E}_{T-1}l_{nT}, \mathbb{E}_{T-1}\mu_{niT})$.

Using the fact that $\mathbb{E}_{T-2}\hat{l}_{T-1} = 0$, we can sequentially solve the system. We obtain $\{\mathbb{E}_{T-2}\lambda_{nit}, \mathbb{E}_{T-2}P_{nt}, \mathbb{E}_{T-2}w_{nt}, \mathbb{E}_{T-2}l_{nt}, \mathbb{E}_{T-2}\mu_{nit}\}_{i=1,n=1,t=T-1}^{N,N,T}$.

iv Append the output from (iii) with $\tilde{\mu}_{niT-2}$, agents' actual migration decision in period $T-2$. Together, $\tilde{\mu}_{niT-2}$ and the output from (iii) constitute the solution to the agents' problem in period T-2.

v Repeat (iii) and (iv) recursively backward until the first period is reached.

**Extension for when the data end at $T' < T$.** In the above discussion we have assumed that the researcher observed the actual allocation until the end of model $T$. Suppose the researcher only observed the outcome up until $T' < T$ (more precisely, the researcher observed labor market outcomes in $T'$, but only migration decisions in $T'-1$). We proceed by first constructing the belief path of agents when they make a decision in period $T'-1$ using a time-differenced version of (2)-(7) until the last period of the model, then start from Step (iv) of the algorithm described in Section 2.5 and iterate backward according to the algorithm.

The belief path of agents at $T'-1$ period solves the following set of equations

$$\begin{aligned}
&\text{For } \ t \geq T', \\
\mathbb{E}_{T'-1}\hat{v}_{nt} &= \frac{\partial U(\bar{w}_{nT'},\bar{P}_{nT'})}{\partial \log \bar{w}_{nT'}}\mathbb{E}_{T'-1}\hat{w}_{nt} + \frac{\partial U(\bar{w}_{nT'},\bar{P}_{nT'})}{\partial \log \bar{P}_{nT'}}\mathbb{E}_{T'-1}\hat{P}_{nt} + \beta\sum_{i=1}^{N}\bar{\mu}_{niT'-1}\mathbb{E}_{T'-1}\hat{v}_{it+1} \\
\mathbb{E}_{T'-1}\hat{\mu}_{nit} &= \frac{\beta}{\nu}\sum_{k=1}^{N}\Big(\mathbb{1}(k=i)-\bar{\mu}_{nkT'-1}\Big)\mathbb{E}_{T'-1}\hat{v}_{kt+1} \\
\mathbb{E}_{T'-1}\hat{l}_{nt} &= \sum_{i}^{N}\frac{\bar{\mu}_{inT'-1}\bar{l}_{iT'-1}}{\bar{l}_{nT'}}\Big(\mathbb{E}_{T'-1}\hat{\mu}_{int-1}+\mathbb{E}_{T'-1}\hat{l}_{it-1}\Big) \\
\mathbb{E}_{T'-1}\hat{\lambda}_{nit} &= -\theta(\mathbb{E}_{T'-1}\hat{w}_{it}+\mathbb{E}_{T'-1}\hat{\kappa}_{int}-\mathbb{E}_{T'-1}\hat{P}_{nt})+\mathbb{E}_{T'-1}\hat{z}_{it} \\
\mathbb{E}_{T'-1}\hat{P}_{nt} &= \sum_{i}\bar{\lambda}_{niT'}\Big(\mathbb{E}_{T'-1}\hat{w}_{it}+\mathbb{E}_{T'-1}\hat{\kappa}_{int}-\frac{1}{\theta}\mathbb{E}_{T'-1}\hat{z}_{it}\Big) \\
\mathbb{E}_{T'-1}\hat{w}_{nt}+\mathbb{E}_{T'-1}\hat{l}_{nt} &= \sum_{i}\frac{\bar{\lambda}_{inT'}\bar{w}_{iT'}\bar{l}_{iT'}}{\bar{w}_{nT'}\bar{l}_{nT'}}\Big(\hat{\lambda}_{int}+\hat{w}_{it}+\hat{l}_{it}\Big),
\end{aligned} \tag{A.18}$$

where $\mathbb{E}_{T'-1}\hat{x}_t \equiv \mathbb{E}_{T'-1}\log(x_t)-\log(x_{T'})$ for $x \in \{w,P,l,\lambda,\kappa,z\}$ and $\hat{\mu}_t \equiv \mathbb{E}_{T'-1}\log(\mu_t)-\log(\mu_{T'-1})$.

**Recovering welfare from the actual and belief paths.** Let $v^t_{nt'}$ be the expected value of location $n$ in period $t$ under the belief from period $t'$. Let $\tilde{v}_{nt}$ be the actual value of location $n$ at period $t$, which is defined ex-post to be the average value of a location and takes into account the fact that the people in each location make the best decision according to their beliefs but not the best decision ex-post.

Denote by $T$ the last period of the model. Without loss of generality, let period $\tilde{T}$ the first period people become perfect foresight, so $v^t_{n\tilde{T}} = v_{nt}$ for $t \geq \tilde{T}$. Then $\tilde{v}^{\tilde{T}}_{n\tilde{T}-1}$ and the actual welfare $\tilde{v}_{nt}$ for $t \geq \tilde{T}$ can be computed as in Caliendo et al. (2019):

$$\begin{aligned}
\tilde{v}_{nt} &= \sum_{s=t}^{T}\beta^s U(\tilde{c}_{ns}) - \nu\sum_{s=t}^{T}\beta^s\log\tilde{\mu}_{nis} \\
v^{\tilde{T}}_{n\tilde{T}-1} &= \sum_{s=\tilde{T}}^{T}\beta^s U(\mathbb{E}_{\tilde{T}-1}c_{ns}) - \nu\sum_{s=\tilde{T}}^{T}\beta^s\log\mathbb{E}_{\tilde{T}-1}\mu_{nis},
\end{aligned} \tag{A.19}$$

where $\tilde{c}_{ns} \equiv \frac{\tilde{w}_{ns}}{\tilde{P}_{ns}}$ is the realized real wage at $s$, and $\mathbb{E}_{\tilde{T}-1}c_{ns}$ the expected value according to belief at period $T-1$. By evaluating $U$ and $log$ functions at the expected outcomes, equation (A.19) uses the certainty equivalence result.

Now, consider the actual value of location $n$ at $\tilde{T}-1$ period, $\tilde{v}_{n\tilde{T}-1}$:

$$\begin{aligned}
\tilde{v}_{n\tilde{T}-1} &= U(\tilde{c}_{n\tilde{T}-1}) + \mathbb{E}_{\tilde{T}-1}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta\tilde{v}_{j\tilde{T}} - m_{nj\tilde{T}-1}+\epsilon_j] \\
&= U(\tilde{c}_{n\tilde{T}-1}) + \mathbb{E}_{\tilde{T}-1}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta v^{\tilde{T}}_{j\tilde{T}-1} - m_{nj\tilde{T}-1}+\epsilon_j] + \mathbb{E}_{\tilde{T}-1}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta\tilde{v}_{j\tilde{T}} - \beta v^{\tilde{T}}_{j\tilde{T}-1}] \\
&= U(\tilde{c}_{n\tilde{T}-1}) + [\beta v^{\tilde{T}}_{n\tilde{T}-1} - \nu\log(\tilde{\mu}_{nn\tilde{T}-1})] + \sum_j \tilde{\mu}_{nj\tilde{T}-1}[\beta\tilde{v}_{j\tilde{T}} - \beta v^{\tilde{T}}_{j\tilde{T}-1}],
\end{aligned} \tag{A.20}$$

where $v^{\tilde{T}}_{j\tilde{T}-1}$ can be calculated using equation (A.19) applying the belief at $\tilde{T}-1$. Next, consider $\tilde{v}_{n\tilde{T}-2}$:

$$\begin{aligned}
\tilde{v}_{n\tilde{T}-2} &= U(\tilde{c}_{n\tilde{T}-2}) + \mathbb{E}_{\tilde{T}-2}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta v_{j\tilde{T}-1} - m_{nj\tilde{T}-2} + \epsilon_j] \\
&= U(\tilde{c}_{n\tilde{T}-2}) + \mathbb{E}_{\tilde{T}-2}\sum_j \mathbf{1}_{\text{j is chosen}}[\beta v^{T-1}_{j\tilde{T}-2} - m_{nj\tilde{T}-2} + \epsilon_j] + \mathbb{E}_{\tilde{T}-2}\sum_j \mathbf{1}_j[\beta\tilde{v}_{j\tilde{T}-1} - \beta v^{\tilde{T}-1}_{j\tilde{T}-2}] \\
&= U(\tilde{c}_{n\tilde{T}-2}) + [\beta v^{\tilde{T}-1}_{n\tilde{T}-2} - \nu\log(\tilde{\mu}_{nn\tilde{T}-2})] + \sum_j \tilde{\mu}_{nj\tilde{T}-2}[\beta\tilde{v}_{j\tilde{T}-1} - \beta v^{\tilde{T}-1}_{j\tilde{T}-2}],
\end{aligned} \tag{A.21}$$

where $v^{\tilde{T}-1}_{j\tilde{T}-2}$ can be calculated using equation (A.19) applying the belief at $\tilde{T}-2$.

Recursively applying this backward, we obtain $\tilde{v}_{n1} = U(\tilde{c}_{n1}) + [\beta v^2_{n1} - \nu\log(\tilde{\mu}_{nn1})] + \sum_j \tilde{\mu}_{nj1}[\beta\tilde{v}_{j1} - \beta v^2_{j1}]$.

## A.3 Proof of Proposition 3

A second-order approximation of equations (2)-(7) are (see Appendix A.6.2 for derivations):

$$
\begin{aligned}
\mathbb{E}_t\hat{v}_{nt+1} &= \frac{\partial U(\bar{w}_{nt+1},\bar{P}_{nt+1})}{\partial\log(\bar{w}_{nt+1})}\mathbb{E}_t\hat{w}_{nt+1} + \frac{\partial U(\bar{w}_{nt+1},\bar{P}_{nt+1})}{\partial\log(\bar{P}_{nt+1})}\mathbb{E}_t\hat{P}_{nt+1} \\
&\quad + \frac{1}{2}\frac{\partial^2 U(\bar{w}_{nt+1},\bar{P}_{nt+1})}{\partial\log(\bar{w}_{nt+1})^2}\mathbb{E}_t\hat{w}^2_{nt+1} + \frac{1}{2}\frac{\partial^2 U(\bar{w}_{nt+1},\bar{P}_{nt+1})}{\log(\partial\bar{P}_{nt+1})^2}\mathbb{E}_t\hat{P}^2_{nt+1} + \frac{\partial^2 U(\bar{w}_{nt+1},\bar{P}_{nt+1})}{\partial\log(\bar{w}_{nt+1})\partial\log(\bar{P}_{nt+1})}\mathbb{E}_t\hat{w}_{nt+1}\hat{P}_{nt+1} \\
&\quad + \sum_k \beta\bar{\mu}_{nkt+1}\mathbb{E}_t\hat{v}_{kt+2} + \frac{1}{2}\sum_m\sum_k\frac{\beta^2}{\nu}\bar{\mu}_{nmt+1}\Big(\mathbb{I}(k=m)-\bar{\mu}_{nkt+1}\Big)\mathbb{E}_t\hat{v}_{mt+2}\hat{v}_{kt+2} \\
\mathbb{E}_t\hat{\mu}_{nit} &= \sum_k\frac{\beta}{\nu}\Big(\mathbb{I}(k=i)-\bar{\mu}_{nkt}\Big)\mathbb{E}_t\hat{v}_{kt+1} \\
&\quad + \frac{1}{2}\sum_{k,m}(\frac{\beta}{\nu})^2\cdot\Big[\mathbb{I}(k=i)\bar{\mu}_{nit}(\mathbb{I}(m=i)-\bar{\mu}_{nmt})-\bar{\mu}_{nkt}\bar{\mu}_{nit}[\mathbb{I}(m=i)-2\bar{\mu}_{nmt}+\mathbb{I}(m=k)]\Big]\mathbb{E}_t\hat{v}_{kt+1}\hat{v}_{mt+1} \\
\mathbb{E}_t\hat{l}_{nt+1} &= \sum_{k=1}^N\frac{\beta}{\nu}\sum_{i=1}^N\bar{\psi}_{int+1}\Big(\mathbb{I}(k=n)-\bar{\mu}_{ikt}\Big)\mathbb{E}_t\hat{v}_{kt+1} + \sum_{k=1}^N\bar{\psi}_{knt+1}\mathbb{E}_t\hat{l}_{kt} \\
&\quad + \frac{1}{2}\sum_{k,m}(\frac{\beta}{\nu})^2\cdot\Big[-\sum_{i=1}^N\bar{\mu}_{ikt}\bar{\psi}_{int+1}[\mathbb{I}(m=n)+\mathbb{I}(m=k)-2\bar{\mu}_{imt}]+\mathbb{I}(k=n)\sum_i\bar{\psi}_{int+1}(\mathbb{I}(m=n)-\bar{\mu}_{imt}) \\
&\quad - [\sum_{i=1}^N\bar{\psi}_{int+1}(\mathbb{I}(k=n)-\bar{\mu}_{ikt})]\cdot[\sum_i\bar{\psi}_{int+1}(\mathbb{I}(m=n)-\bar{\mu}_{imt})]\Big]\mathbb{E}_t\hat{v}_{kt+1}\hat{v}_{mt+1} \\
&\quad + \frac{1}{2}\sum_{k,m}[\mathbf{1}(k=m)\bar{\psi}_{mnt+1}-\bar{\psi}_{knt+1}\bar{\psi}_{mnt+1}]\,\mathbb{E}_t\hat{l}_{kt}\hat{l}_{mt} \\
&\quad + \sum_{k,m}\left[\frac{\beta}{\nu}\bar{\psi}_{mnt+1}(\mathbb{I}(k=n)-\bar{\mu}_{mkt})-\frac{\beta}{\nu}\bar{\psi}_{mnt+1}[\sum_i\bar{\psi}_{int+1}(\mathbb{I}(k=n)-\bar{\mu}_{ikt})]\right]\mathbb{E}_t\hat{v}_{kt+1}\hat{l}_{mt}, \\
\mathbb{E}_t\hat{\lambda}_{nit} &= -\theta(\mathbb{E}_t\hat{w}_{it}+\mathbb{E}_t\hat{\kappa}_{int}-\mathbb{E}_t\hat{P}_{nt})+\mathbb{E}_t\hat{z}_{it} \\
\mathbb{E}_t\hat{P}_{nt} &= \sum_{i=1}^N\bar{\lambda}_{nit}\mathbb{E}_t\hat{w}_{it}+\sum_{i=1}^N\bar{\lambda}_{nit}\cdot\mathbb{E}_t\hat{w}_{it}+\sum_{i=1}^N\Big(-\frac{1}{\theta}\Big)\bar{\lambda}_{nit}\mathbb{E}_t\hat{z}_{it} \\
&\quad + \frac{1}{2}\sum_{i,m}\theta\Big(-\mathbf{I}(i=m)\bar{\lambda}_{nit}+\bar{\lambda}_{nit}\bar{\lambda}_{nmt}\Big)\cdot\Big(\mathbb{E}_t\hat{w}_{it}\hat{w}_{mt}+\mathbb{E}_t\hat{\kappa}_{nit}\hat{\kappa}_{nmt}+\Big(\frac{1}{\theta}\Big)^2\mathbb{E}_t\hat{z}_{it}\hat{z}_{mt}\Big) \\
&\quad + \sum_{i,m}\theta\Big(-\mathbf{I}(i=m)\bar{\lambda}_{nit}+\bar{\lambda}_{nit}\bar{\lambda}_{nmt}\Big)\Big(\mathbb{E}_t\hat{w}_{it}\hat{\kappa}_{nmt}-\Big(\frac{1}{\theta}\Big)\mathbb{E}_t\hat{w}_{it}\hat{z}_{mt}-\Big(\frac{1}{\theta}\Big)\mathbb{E}_t\hat{\kappa}_{nit}\hat{z}_{mt}\Big) \\
\mathbb{E}_t\hat{w}_{nt}+\mathbb{E}_t\hat{l}_{nt} &= \sum_{i=1}^N\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{\bar{w}_{nt}\bar{l}_{nt}}\Big(\mathbb{E}_t\hat{\lambda}_{int}+\mathbb{E}_t\hat{w}_{it}+\mathbb{E}_t\hat{l}_{it}\Big) \\
&\quad + \frac{1}{2}\sum_{i,m}\Big(1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})}-\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}\Big)\Big(\mathbb{E}_t\hat{w}_{it}\hat{w}_{mt}+\mathbb{E}_t\hat{\lambda}_{int}\hat{\lambda}_{mnt}+\mathbb{E}_t\hat{l}_{it}\hat{l}_{mt}\Big) \\
&\quad + \sum_{i,m}\Big(1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})}-\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}\Big)\Big(\mathbb{E}_t\hat{\lambda}_{int}\hat{w}_{mt}+\mathbb{E}_t\hat{\lambda}_{int}\hat{l}_{mt}+\mathbb{E}_t\hat{w}_{it}\hat{l}_{mt}\Big),
\end{aligned}
\tag{A.22}
$$

where $\bar{\psi}_{int+1} = \frac{\bar{\mu}_{int}\bar{l}_{it}}{\bar{l}_{nt+1}}$.

The equilibrium objects are $\Big\{\mathbb{E}_t\hat{v}_{it'},\mathbb{E}_t\hat{\mu}_{nit'},\mathbb{E}_t\hat{l}_{nt'},\mathbb{E}_t\hat{w}_{it'},\mathbb{E}_t\hat{P}_{it'}\Big\}_{i=1,n=1,t'=t}^{N,N,T}$ and the additional inputs compared to Proposition 1 are the covariance of these variables, e.g., $\{\mathbb{E}_t\hat{v}_{it'}\hat{v}_{nt'},\mathbb{E}_t\hat{w}_{it'}\hat{w}_{nt'},\mathbb{E}_t\hat{P}_{it'}\hat{P}_{nt'}\}$

We propose the following algorithm using simulation which gives second-order accuracy.

1. Start from period $t=t_0$; simulate S paths of $z_{t_0}^T$.

2. Solve the first-order approximation for each $s=1,..,S$, denote the solution $\hat{x}_{t'}(s)$ for all $t'\geq t_0$, where $\hat{x}\in\{\hat{w}_{nt'},\hat{P}_{nt'},\hat{v}_{t'+1},etc.\}$

3. Approximate $\mathbb{E}_{t_0}\hat{x}_{t'}\hat{y}_{t'} \approx \frac{1}{S}\sum_{s=1}^{S}\hat{x}_{t'}(s)\hat{y}_{t'}(s)$; plug into the second order approximation equations (equations (A.22)) and solve for the remaining terms, which gives the decision in period $t_0$.

4. Move the economy to $t = t_0 + 1$, repeat the above process until $t = T$.

## A.4 Proof of Proposition 4

In this section, we prove Proposition 4 under each assumption separately. In each case, we provide a brief discussion on the numerical algorithm.

### A.4.1 Assumption 1

We denote the two groups of agents by $A$ and $B$, with their beliefs given by $f^g(z_{t+1}|z^t)$, $g \in A, B$. Let $\mathbb{E}_t^g z_{t+1}$ the expected productivity for period $t+1$ under type $g \in \{A, B\}$'s belief at $t$, that is, $\mathbb{E}_t^g z_{t+1} = \int_{z_{t+1}} z_{t+1} f^g(z_{t+1}|z^t)dz_{t+1}$. Then, the expected value of location $n$ for type $g$ agent is

$$\begin{aligned}
\mathbb{E}^g[v_{nt+1}^g(z^{t+1})|z^t] &= \int_{z_{t+1}} U(c_{nt}^g(z^t)) + \nu \log\left(\sum_{i=1}^{N}\exp\left(\beta\mathbb{E}_t^g[v_{it+2}^g(z^{t+1})|z^t] - m_{nit}\right)^{1/\nu}\right) f^g(z_{t+1}|z^t)dz_{t+1} \\
&\equiv \mathbb{E}_t^g v_{nt+1}^g
\end{aligned} \tag{A.23}$$

Similarly, the migration decision for type $g$ is given by

$$\mu_{nit}^g(z^t) = \frac{\exp\left(\beta\mathbb{E}_t^g v_{it+1}^g - m_{nit}\right)^{1/\nu}}{\sum_{h=1}^{N}\exp\left(\beta\mathbb{E}_t^g v_{ht+1}^g - m_{nht}\right)^{1/\nu}}. \tag{A.24}$$

While this equation has the same form as equation (2), wage $w_{n,t}(z^t)$ and price $P_{n,t}(z^t)$ are determined by the decisions of both types given $(z^t)$, which are given by the solution to the following system of equations:

$$\begin{aligned}
\lambda_{nit}(z^t) &= z_{it}\left(\frac{w_{it}(z^t)\kappa_{nit}}{P_{nt}(z^t)}\right)^{-\theta} \\
P_{nt}(z^t) &= \left[\sum_{i=1}^{N} z_{it}\left(w_{it}(z^t)\kappa_{nit}\right)^{-\theta}\right]^{-1/\theta} \\
w_{nt}(z^t)l_{nt}(z^t) &= \sum_{i=1}^{N}\lambda_{int}(z^t)w_{it}(z^t)l_{it}(z^t) \\
l_{nt+1}(z^t) &= \sum_{g=A,B} l_{nt+1}^g(z^t) \\
l_{nt+1}^g(z^t) &= \sum_{n}\mu_{mnt}^g(z^t)l_{mt}^g(z^{t-1}), \text{ for } g \in \{A, B\}
\end{aligned} \tag{A.25}$$

We first consider the first-period problem of agent $A$. Following the approach of Proposition 1, we linearize around a homogeneous belief perfect foresight equilibrium and sum across the equations for all the possible states in each period weighted by $f^A(z_{t+1}|z^t)$. For example, type A agents' expectation on future productivity given history $z^t$ is defined by:

$$
\begin{aligned}
\mathbb{E}_t^A z_{t+1} &\equiv \int_{z_{t+1}} z_{n,t+1} f^A(z_{t+1}|z^t) dz_{t+1} \\
\mathbb{E}_t^A z_{t+2} &\equiv \int_{z_{t+2}} z_{t+2} f^A(z_{t+2}|z^t) dz_{t+2} \\
&= \int_{z_{t+2}} z_{t+2} f^A(z_{t+2}|z^{t+1}) f^A(z_{t+1}|z^t) dz_{t+2} \qquad \text{(A.26)}
\end{aligned}
$$

This gives us the following system of linear expectation equations for type $A$ at period $t$.

$$
\begin{aligned}
\mathbb{E}_t^A \hat{v}_{nt+1}^A &= \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{w}_{nt+1})} \mathbb{E}_t^A \hat{w}_{nt+1} + \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{P}_{nt+1})} \mathbb{E}_t^A \hat{P}_{nt+1} \\
&\quad + \beta \sum_i \bar{\mu}_{ni,t+1} \mathbb{E}_t^A \hat{v}_{nt+2}^A, \qquad \text{(A.27)} \\
\hat{\mu}_{ni,t}^A &= \frac{\beta}{\nu} \mathbb{E}_t^A \hat{v}_{it+1}^A - \frac{\beta}{\nu} \sum_{m=1}^N \bar{\mu}_{nm,t} \beta \mathbb{E}_t^A \hat{v}_{mt+1}^A, \qquad \text{(A.28)} \\
\hat{l}_{nt+1}^A &= \sum_{i=1}^N \bar{\psi}_{int+1}^A \left( \hat{\mu}_{int}^A + \hat{l}_{it}^A \right), \qquad \text{(A.29)} \\
\hat{l}_{nt+1}^B &= \sum_{i=1}^N \bar{\psi}_{int+1}^B \left( \hat{\mu}_{int}^B + \hat{l}_{it}^B \right), \qquad \text{(A.30)} \\
\hat{l}_{nt+1} &= \frac{\bar{l}_{nt+1}^A}{\bar{l}_{nt+1}} \hat{l}_{nt+1}^A + \frac{\bar{l}_{nt+1}^B}{\bar{l}_{nt+1}} \hat{l}_{nt+1}^B, \qquad \text{(A.31)} \\
\mathbb{E}_t^A \hat{\lambda}_{nit+1} &= -\theta \left( \mathbb{E}_t^A \hat{w}_{it+1} - \mathbb{E}_t^A \hat{P}_{nt+1} + \mathbb{E}_t^A \hat{\kappa}_{ni,t+1} \right) + \mathbb{E}_t^A \hat{z}_{it+1}, \qquad \text{(A.32)} \\
\mathbb{E}_t^A \hat{P}_{nt+1} &= \sum_i \bar{\lambda}_{nit+1} \Big( \mathbb{E}_t^A \hat{w}_{it+1} + \mathbb{E}_t^A \hat{\kappa}_{ni,t+1} - \frac{1}{\theta} \mathbb{E}_t^A \hat{z}_{it+1} \Big), \qquad \text{(A.33)} \\
\mathbb{E}_t^A \hat{w}_{nt+1} + \mathbb{E}_t^A \hat{l}_{nt+1} &= \sum_i \frac{\bar{\lambda}_{int+1} \bar{w}_{nt+1} \bar{l}_{nt+1}}{\bar{w}_{it+1} \bar{l}_{it+1}} \Big( \mathbb{E}_t^A \hat{\lambda}_{int+1} + \mathbb{E}_t^A \hat{w}_{it+1} + \mathbb{E}_t^A \hat{l}_{it+1} \Big) \qquad \text{(A.34)}
\end{aligned}
$$

and for $t' \geq t+1$,

$$
\begin{aligned}
\mathbb{E}_t^A \hat{v}_{nt'+1}^A &= \frac{\partial U(\bar{w}_{nt'+1}, \bar{P}_{nt'+1})}{\partial \log(\bar{w}_{nt'+1})} \mathbb{E}_t^A \hat{w}_{nt'+1} + \frac{\partial U(\bar{w}_{nt'+1}, \bar{P}_{nt'+1})}{\partial \log(\bar{P}_{nt'+1})} \mathbb{E}_t^A \hat{P}_{nt'+1} \\
&\quad + \beta \sum_i \bar{\mu}_{ni,t'+1} \mathbb{E}_t^A \hat{v}_{nt'+2}^A, && \text{(A.35)} \\
\mathbb{E}_t^A \hat{\mu}_{ni,t'}^A &= \frac{\beta}{\nu} \mathbb{E}_t^A \hat{v}_{it'+1}^A - \frac{\beta}{\nu} \sum_{m=1}^N \bar{\mu}_{nm,t'} \beta \mathbb{E}_t^A \hat{v}_{mt'+1}^A, && \text{(A.36)} \\
\mathbb{E}_t^A \hat{l}_{nt'+1}^A &= \sum_{i=1}^N \bar{\psi}_{int'+1}^A \left( \mathbb{E}_t^A \hat{\mu}_{int'}^A + \mathbb{E}_t^A \hat{l}_{it'}^A \right), && \text{(A.37)} \\
\mathbb{E}_t^A \hat{l}_{nt'+1}^B &= \sum_{i=1}^N \bar{\psi}_{int'+1}^B \left( \mathbb{E}_t^A \hat{\mu}_{int'}^B + \mathbb{E}_t^A \hat{l}_{it'}^B \right), && \text{(A.38)} \\
\mathbb{E}_t^A \hat{l}_{nt'+1} &= \frac{\bar{l}_{nt'+1}^A}{\bar{l}_{nt'+1}} \mathbb{E}_t^A \hat{l}_{nt'+1}^A + \frac{\bar{l}_{nt'+1}^B}{\bar{l}_{nt'+1}} \mathbb{E}_t^A \hat{l}_{nt'+1}^B, && \text{(A.39)} \\
\mathbb{E}_t^A \hat{\lambda}_{ni,t'} &= -\theta \left( \mathbb{E}_t^A \hat{w}_{it'} - \mathbb{E}_t^A \hat{P}_{nt'} + \mathbb{E}_t^A \hat{\kappa}_{nit'} \right) + \mathbb{E}_t^A \hat{z}_{it'}, && \text{(A.40)} \\
\mathbb{E}_t^A \hat{P}_{nt'} &= \sum_i \bar{\lambda}_{ni,t'} \left( \mathbb{E}_t^A \hat{w}_{it'} + \mathbb{E}_t^A \hat{\kappa}_{nit'} - \frac{1}{\theta} \mathbb{E}_t^A \hat{z}_{it'} \right), && \text{(A.41)} \\
\mathbb{E}_t^A \hat{w}_{nt'} + \mathbb{E}_t^A \hat{l}_{nt'} &= \sum_i \frac{\bar{\lambda}_{int'} \bar{w}_{nt'} \bar{l}_{nt'}}{\bar{w}_{it'} \bar{l}_{it'}} \left( \mathbb{E}_t^A \hat{\lambda}_{int'} + \mathbb{E}_t^A \hat{w}_{it'} + \mathbb{E}_t^A \hat{l}_{it'} \right) && \text{(A.42)}
\end{aligned}
$$

where $\bar{\psi}_{int+1}^g \equiv \frac{\bar{\mu}_{int}^g \bar{l}_{it}^g}{\bar{l}_{nt+1}^g}$ for $g \in \{A, B\}$ and $\bar{l}_{nt+1} = \bar{l}_{nt+1}^A + \bar{l}_{nt+1}^B$. These expectational equations, together with the equation for the current period migration decision, constitute a solution to the period $t$ problem. As shown in equation (A.38), to solve for A's decision in period $t$, we need $A$'s belief about B's future migration decisions, $\mathbb{E}_t^A \mu_{nit'}^B$, for all $t' > t$:

$$
\mathbb{E}_t^A \hat{\mu}_{ni,t'}^B = \frac{\beta}{\nu} \mathbb{E}_t^A \mathbb{E}_{t'}^B \hat{v}_{it'+1}^B - \frac{\beta}{\nu} \sum_{m=1}^N \bar{\mu}_{nm,t'} \beta \mathbb{E}_t^A \mathbb{E}_{t'}^B \hat{v}_{mt'+1}^B, \tag{A.43}
$$

Type $B$'s migration decision for $t'$ can be characterized by the solution to the linearized equations for all the possible states in each period weighted by B's belief, $f^B(z_{t''}|z^{t'})$, $t'' \geq t'$.

$$\begin{aligned}
\mathbb{E}^B_{t'}\hat{v}^B_{nt''+1} &= \frac{\partial U(\bar{w}_{nt''+1},\bar{P}_{nt''+1})}{\partial\log(\bar{w}_{nt''+1})}\mathbb{E}^B_{t'}\hat{w}_{nt''+1} + \frac{\partial U(\bar{w}_{nt''+1},\bar{P}_{nt''+1})}{\partial\log(\bar{P}_{nt''+1})}\mathbb{E}^B_{t'}\hat{P}_{nt''+1} \\
&\quad +\beta\sum_i \bar{\mu}_{ni,t''+1}\mathbb{E}^B_{t'}\hat{v}^B_{nt''+2}, \qquad \text{(A.44)} \\
\mathbb{E}^B_{t'}\hat{\mu}^B_{ni,t''} &= \frac{\beta}{\nu}\mathbb{E}^B_{t'}\hat{v}^B_{it''+1} - \frac{\beta}{\nu}\sum_{m=1}^{N}\bar{\mu}_{nm,t''}\beta\mathbb{E}^B_{t'}\hat{v}^B_{mt''+1}, \qquad \text{(A.45)} \\
\mathbb{E}^B_{t'}\hat{l}^A_{nt''+1} &= \sum_{i=1}^{N}\bar{\psi}^A_{int''+1}\left(\mathbb{E}^B_{t'}\hat{\mu}^A_{int''} + \mathbb{E}^B_{t'}\hat{l}^A_{it''}\right), \qquad \text{(A.46)} \\
\mathbb{E}^B_{t'}\hat{l}^B_{nt''+1} &= \sum_{i=1}^{N}\bar{\psi}^B_{int''+1}\left(\mathbb{E}^B_{t'}\hat{\mu}^B_{int''} + \mathbb{E}^B_{t'}\hat{l}^B_{it''}\right), \qquad \text{(A.47)} \\
\mathbb{E}^B_{t'}\hat{l}_{nt''+1} &= \frac{\bar{l}^A_{nt''+1}}{\bar{l}_{nt''+1}}\mathbb{E}^B_{t'}\hat{l}^A_{nt''+1} + \frac{\bar{l}^B_{nt''+1}}{\bar{l}_{nt''+1}}\mathbb{E}^B_{t'}\hat{l}^B_{nt''+1}, \qquad \text{(A.48)} \\
\mathbb{E}^B_{t'}\hat{\lambda}_{ni,t''} &= -\theta\left(\mathbb{E}^B_{t'}\hat{w}_{it'} - \mathbb{E}^B_{t'}\hat{P}_{nt''} + \mathbb{E}^B_{t'}\hat{\kappa}_{nit''}\right) + \mathbb{E}^B_{t'}\hat{z}_{it''}, \qquad \text{(A.49)} \\
\mathbb{E}^B_{t'}\hat{P}_{nt''} &= \sum_i \bar{\lambda}_{nit''}\left(\mathbb{E}^B_{t'}\hat{w}_{it''} + \mathbb{E}^B_{t'}\hat{\kappa}_{nit''} - \frac{1}{\theta}\mathbb{E}^B_{t'}\hat{z}_{it''}\right), \qquad \text{(A.50)} \\
\mathbb{E}^B_{t'}\hat{w}_{nt''} + \mathbb{E}^B_{t'}\hat{l}_{nt''} &= \sum_i \frac{\bar{\lambda}_{int''}\bar{w}_{nt''}\bar{l}_{nt''}}{\bar{w}_{it''}\bar{l}_{it''}}\left(\mathbb{E}^B_{t'}\hat{\lambda}_{int''} + \mathbb{E}^B_{t'}\hat{w}_{it''} + \mathbb{E}^B_{t'}\hat{l}_{it''}\right) \qquad \text{(A.51)}
\end{aligned}$$

Note that this equation holds for any $t'' = t', ..., T$. *In the eyes of type $A$ agents at period $t$*, type $B$'s migration decision at $t''$ can characterized by

$$\begin{aligned}
\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{v}^B_{nt''+1} &= \frac{\partial U(\bar{w}_{nt''+1},\bar{P}_{nt''+1})}{\partial\log(\bar{w}_{nt''+1})}\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{w}_{nt''+1} + \frac{\partial U(\bar{w}_{nt''+1},\bar{P}_{nt''+1})}{\partial\log(\bar{P}_{nt''+1})}\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{P}_{nt''+1} \\
&\quad +\beta\sum_i \bar{\mu}_{ni,t''+1}\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{v}^B_{nt''+2}, \qquad \text{(A.52)} \\
\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{\mu}^B_{ni,t''} &= \frac{\beta}{\nu}\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{v}^B_{it'+1} - \frac{\beta}{\nu}\sum_{m=1}^{N}\bar{\mu}_{nm,t''}\beta\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{v}^B_{mt'+1},, \qquad \text{(A.53)} \\
\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}^A_{nt''+1} &= \sum_{i=1}^{N}\bar{\psi}^A_{int''+1}\left(\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{\mu}^A_{int''} + \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}^A_{it'}\right), \qquad \text{(A.54)} \\
\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}^B_{nt''+1} &= \sum_{i=1}^{N}\bar{\psi}^B_{int''+1}\left(\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{\mu}^B_{int''} + \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}^B_{it'}\right), \qquad \text{(A.55)} \\
\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}_{nt''+1} &= \frac{\bar{l}^A_{nt''+1}}{\bar{l}_{nt''+1}}\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}^A_{nt''+1} + \frac{\bar{l}^B_{nt''+1}}{\bar{l}_{nt''+1}}\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}^B_{nt''+1}, \qquad \text{(A.56)} \\
\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{\lambda}_{ni,t''} &= -\theta\left(\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{w}_{it''} - \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{P}_{nt''} + \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{\kappa}_{nit''}\right) + \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{z}_{it''}, \qquad \text{(A.57)} \\
\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{P}_{nt''} &= \sum_i \bar{\lambda}_{nit''}\left(\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{w}_{it''} + \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{\kappa}_{nit''} - \frac{1}{\theta}\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{z}_{it''}\right), \qquad \text{(A.58)} \\
\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{w}_{nt''} + \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}_{nt''} &= \sum_i \frac{\bar{\lambda}_{int''}\bar{w}_{nt''}\bar{l}_{nt''}}{\bar{w}_{it''}\bar{l}_{it''}}\left(\mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{\lambda}_{int''} + \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{w}_{it''} + \mathbb{E}^A_t\mathbb{E}^B_{t'}\hat{l}_{it''}\right) \qquad \text{(A.59)}
\end{aligned}$$

This process continues as type $B$'s belief on type $A$'s future migration decision, $\mathbb{E}^B_{t'}\hat{\mu}^A_{ni,t''}$, matters in type $B$'s migration decision at $t'$. The implication of this process is that the higher order beliefs about future fundamentals: $\mathbb{E}^A_1\hat{z}_t$, $\mathbb{E}^A_1\mathbb{E}^B_2\hat{z}_t$, $\mathbb{E}^A_1\mathbb{E}^B_2\mathbb{E}^A_3\hat{z}_t$, $\mathbb{E}^A_1\mathbb{E}^B_2\mathbb{E}^A_3\mathbb{E}^B_4\hat{z}_t$, ..... are relevant for the decision of agent $A$ in the first period. Solving this problem, therefore, requires calculating all higher-order beliefs about fundamentals and solving the higher-order beliefs about agents' endogenous decisions, which is why additional assumptions are necessary.

An immediate consequence of Assumption 1 is that higher-order beliefs of each type $g \in \{A, B\}$ and $g' \neq g$ can be simplified as

$$\mathbb{E}^g_t\hat{z}_{t''} = \mathbb{E}^g_t\mathbb{E}^{g'}_{t'}\hat{z}_{t''} \text{ for } t' \geq t+1 \text{ and } t'' \geq t'. \tag{A.60}$$

This implies we can simplify the high-order beliefs in (A.52) to (A.59) to the following

$$\begin{aligned}
\mathbb{E}^A_t\hat{v}^B_{nt''+1} &= \frac{\partial U(\bar{w}_{nt''+1}, \bar{P}_{nt''+1})}{\partial\log(\bar{w}_{nt''+1})}\mathbb{E}^A_t\hat{w}_{nt''+1} + \frac{\partial U(\bar{w}_{nt''+1}, \bar{P}_{nt''+1})}{\partial\log(\bar{P}_{nt''+1})}\mathbb{E}^A_t\hat{P}_{nt''+1} \\
&\quad +\beta\sum_i \bar{\mu}_{ni,t''+1}\mathbb{E}^A_t\hat{v}^B_{nt''+2}, & \text{(A.61)}\\
\mathbb{E}^A_t\hat{\mu}^B_{ni,t''} &= \frac{\beta}{\nu}\mathbb{E}^A_t\hat{v}^B_{it''+1} - \frac{\beta}{\nu}\sum_{m=1}^N \bar{\mu}_{nm,t''}\beta\mathbb{E}^A_t\hat{v}^B_{mt''+1}, & \text{(A.62)}\\
\mathbb{E}^A_t\hat{l}^A_{nt''+1} &= \sum_{i=1}^N \bar{\psi}^A_{int''+1}\left(\mathbb{E}^A_t\hat{\mu}^A_{int''} + \mathbb{E}^A_t\hat{l}^A_{it''}\right), & \text{(A.63)}\\
\mathbb{E}^A_t\hat{l}^B_{nt''+1} &= \sum_{i=1}^N \bar{\psi}^B_{int''+1}\left(\mathbb{E}^A_t\hat{\mu}^B_{int''} + \mathbb{E}^A_t\hat{l}^B_{it''}\right), & \text{(A.64)}\\
\mathbb{E}^A_t\hat{l}_{nt''+1} &= \frac{\bar{l}^A_{nt''+1}}{\bar{l}_{nt''+1}}\mathbb{E}^A_t\hat{l}^A_{nt''+1} + \frac{\bar{l}^B_{nt''+1}}{\bar{l}_{nt''+1}}\mathbb{E}^A_t\hat{l}^B_{nt''+1}, & \text{(A.65)}\\
\mathbb{E}^A_t\hat{\lambda}_{ni,t''} &= -\theta\left(\mathbb{E}^A_t\hat{w}_{it''} - \mathbb{E}^A_t\hat{P}_{nt''} + \mathbb{E}^A_t\hat{\kappa}_{nit''}\right) + \mathbb{E}^A_t\hat{z}_{it''}, & \text{(A.66)}\\
\mathbb{E}^A_t\hat{P}_{nt''} &= \sum_i \bar{\lambda}_{nit''}\left(\mathbb{E}^A_t\hat{w}_{it''} + \mathbb{E}^A_t\hat{\kappa}_{nit''} - \frac{1}{\theta}\mathbb{E}^A_t\hat{z}_{it''}\right), & \text{(A.67)}\\
\mathbb{E}^A_t\hat{w}_{nt''} + \mathbb{E}^A_t\hat{l}_{nt''} &= \sum_i \frac{\bar{\lambda}_{int''}\bar{w}_{nt''}\bar{l}_{nt''}}{\bar{w}_{it''}\bar{l}_{it''}}\left(\mathbb{E}^A_t\hat{\lambda}_{int''} + \mathbb{E}^A_t\hat{w}_{it''} + \mathbb{E}^A_t\hat{l}_{it''}\right) & \text{(A.68)}
\end{aligned}$$

Note that equations (A.37)-(A.42) and (A.63)-(A.68) are identical. This together with Assumption 1 implies that *in the eye of type A agents*, both types solve the same problem. In other words, *in the eye of type A agents*, the future value and the migration decision for type $B$ should be the same as that for type $A$. As a result, we can replace $E^A_t\mu^B_{t'}$ by $E^A_t\mu^A_{t'}$ in (A.38). Then the system of linear equations for type $A$ can be solved analogously to Proposition 1 with type $B$'s *current* period migration decision $\hat{\mu}^B_{int}$ as given.

Type $B$'s future decisions, in turn, depends on their beliefs about $A$'s decision in the future. Following the same argument, type $B$ believes the future value and the migration decision for type $A$ should be the same as that for type $B$, and by replacing $\mathbb{E}^B_t\hat{\mu}^A_{int'}$ by $\mathbb{E}^B_t\hat{\mu}^B_{int'}$ into the equations for

$B$ (analogous to equations (A.35)-(A.42)), we obtain the sequence of migration decision for type $B$ given type $A$'s *current* period migration decision $\mathbb{E}_t^B \hat{\mu}_{int}^A$.

Finally, for the solution to be an equilibrium, the belief on the other type's migration decision should be consistent with the actual migration decision.

We solve for the decision in period $t$ and the belief of each type in period $t$ about the future outcomes using the following algorithm:

i Given a guess for $\hat{\mu}_{ni,t}^B$, solve equations (A.27) to (A.42) for A's decision $\hat{\mu}_{ni,t}^A$ and expectation for future outcomes.

ii Given a guess for $\hat{\mu}_{ni,t}^A$, solve equations analogous to (A.27) to (A.42) for B's decision $\hat{\mu}_{ni,t}^B$ and expectation for future outcomes.

iii Iterate on (i) and (ii) until convergence, which gives $\hat{\mu}_{ni,t}^A$ and $\hat{\mu}_{ni,t}^B$.

### A.4.2 Assumption 2

Assumption 2 implies that from third-order beliefs of each type $g \in \{A, B\}$ and $g' \neq g$, the high-order beliefs can be collapsed into the second-order:

$$\mathbb{E}_t^g \hat{z}_{t'''} \neq \mathbb{E}_t^g \mathbb{E}_{t'}^{g'} \hat{z}_{t'''} = \mathbb{E}_t^g \mathbb{E}_{t'}^{g'} \mathbb{E}_{t''}^{g} \hat{z}_{t'''} \text{ for } t' \geq t+1, t'' \geq t'+1, \text{ and } t''' \geq t''. \tag{A.69}$$

We first consider type $A$'s decision. Following the discussion in the case of Assumption 1, Type $A$'s decision at $t$ is given by equations (A.35)-(A.42). Type $B$'s decision at $t'$, which affects type $A$'s decision at $t$, is characterized by equations (A.44)-(A.51). Type $A$'s belief on $B$'s decision is characterized by equations (A.52)-(A.59). Type $B$'s decision at $t'$ depends on type $B$'s belief on type $A$'s future decision, $\mathbb{E}_{t'}^B \hat{\mu}_{ni,t''}^A$, which in turn depends on A's belief in $t''$ about future periods $t''' \geq t''$. Note that A's belief in $t''$ about future outcomes depend on:

$$\mathbb{E}_{t''}^A \hat{\mu}_{ni,t'''}^A = \frac{\beta}{\nu} \mathbb{E}_{t''}^A \hat{v}_{it'''+1}^A - \frac{\beta}{\nu} \sum_{m=1}^{N} \bar{\mu}_{nm,t''} \beta \mathbb{E}_{t''}^A \hat{v}_{mt'''+1}^A, \tag{A.70}$$

where $\mathbb{E}^A_{t''}\hat{v}^A_{it'''+1}$ are characterized by the following system of linear equations:

$$\begin{aligned}
\mathbb{E}^A_{t''}\hat{v}^A_{nt'''+1} &= \frac{\partial U(\bar{w}_{nt'''+1},\bar{P}_{nt'''+1})}{\partial\log(\bar{w}_{nt'''+1})}\mathbb{E}^A_{t''}\hat{w}_{nt'''+1}+\frac{\partial U(\bar{w}_{nt'''+1},\bar{P}_{nt'''+1})}{\partial\log(\bar{P}_{nt'''+1})}\mathbb{E}^A_{t''}\hat{P}_{nt'''+1} \\
&\quad +\beta\sum_i\bar{\mu}_{ni,t'''+1}\mathbb{E}^A_{t''}\hat{v}^A_{nt'''+2}, && \text{(A.71)}\\
\mathbb{E}^A_{t''}\hat{\mu}^A_{ni,t'''} &= \frac{\beta}{\nu}\mathbb{E}^A_{t''}\hat{v}^A_{it'''+1}-\frac{\beta}{\nu}\sum_{m=1}^N\bar{\mu}_{nm,t''}\beta\mathbb{E}^A_{t''}\hat{v}^A_{mt'''+1}, && \text{(A.72)}\\
\mathbb{E}^A_{t''}\hat{l}^A_{nt'''+1} &= \sum_{i=1}^N\bar{\psi}^A_{int'''+1}\left(\mathbb{E}^A_{t''}\hat{\mu}^A_{int'''}+\mathbb{E}^A_{t''}\hat{l}^A_{it'''}\right), && \text{(A.73)}\\
\mathbb{E}^A_{t''}\hat{l}^B_{nt'''+1} &= \sum_{i=1}^N\bar{\psi}^B_{int'''+1}\left(\mathbb{E}^A_{t''}\hat{\mu}^B_{int'''}+\mathbb{E}^A_{t''}\hat{l}^B_{it'''}\right), && \text{(A.74)}\\
\mathbb{E}^A_{t''}\hat{l}_{nt'''+1} &= \frac{\bar{l}^A_{nt'''+1}}{\bar{l}_{nt'''+1}}\mathbb{E}^A_{t''}\hat{l}^A_{nt'''+1}+\frac{\bar{l}^B_{nt'''+1}}{\bar{l}_{nt'''+1}}\mathbb{E}^A_{t''}\hat{l}^B_{nt'''+1}, && \text{(A.75)}\\
\mathbb{E}^A_{t''}\hat{\lambda}_{ni,t'''} &= -\theta\left(\mathbb{E}^A_{t''}\hat{w}_{it'''}-\mathbb{E}^A_{t''}\hat{P}_{nt'''}+\mathbb{E}^A_{t''}\hat{\kappa}_{nit'''}\right)+\mathbb{E}^A_{t''}\hat{z}_{it'''}, && \text{(A.76)}\\
\mathbb{E}^A_{t''}\hat{P}_{nt'''} &= \sum_i\bar{\lambda}_{nit'''}\left(\mathbb{E}^A_{t''}\hat{w}_{it'''}+\mathbb{E}^A_{t''}\hat{\kappa}_{nit'''}-\frac{1}{\theta}\mathbb{E}^A_{t''}\hat{z}_{it'''}\right), && \text{(A.77)}\\
\mathbb{E}^A_{t''}\hat{w}_{nt'''}+\mathbb{E}^A_{t''}\hat{l}_{nt'''} &= \sum_i\frac{\bar{\lambda}_{int'''}\bar{w}_{nt'''}\bar{l}_{nt'''}}{\bar{w}_{it'''}\bar{l}_{it'''}}\left(\mathbb{E}^A_{t''}\hat{\lambda}_{int'''}+\mathbb{E}^A_{t''}\hat{w}_{it''}+\mathbb{E}^A_{t''}\hat{l}_{it'''}\right). && \text{(A.78)}
\end{aligned}$$

Summarizing all above, type $A$'s belief at $t$ on type $B$'s belief at $t'$ on type $A$'s decision from $t'''$

are characterized by the following system of equations

$$\begin{aligned}
\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{v}_{nt'''+1}^A &= \frac{\partial U(\bar{w}_{nt'''+1},\bar{P}_{nt'''+1})}{\partial\log(\bar{w}_{nt'''+1})}\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{w}_{nt'''+1} \\
&\quad+\frac{\partial U(\bar{w}_{nt'''+1},\bar{P}_{nt'''+1})}{\partial\log(\bar{P}_{nt'''+1})}\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{P}_{nt'''+1}+\beta\sum_i\bar{\mu}_{ni,t'''+1}\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{v}_{nt'''+2}^A, && \text{(A.79)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{\mu}_{ni,t'''}^A &= \frac{\beta}{\nu}\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{v}_{it'''+1}^A-\frac{\beta}{\nu}\sum_{m=1}^N\bar{\mu}_{nm,t''}\beta\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{v}_{mt'''+1}^A, && \text{(A.80)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{nt''+1}^A &= \sum_{i=1}^N\bar{\psi}_{int'''+1}^A\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{\mu}_{int'''}^A+\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{it'''}^A\right), && \text{(A.81)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{nt'''+1}^B &= \sum_{i=1}^N\bar{\psi}_{int'''+1}^B\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{\mu}_{int'''}^B+\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{it'''}^B\right), && \text{(A.82)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{nt'''+1} &= \frac{\bar{l}_{nt'''+1}^A}{\bar{l}_{nt'''+1}}\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{nt'''+1}^A+\frac{\bar{l}_{nt'''+1}^B}{\bar{l}_{nt'''+1}}\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{nt'''+1}^B, && \text{(A.83)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{\lambda}_{nit'''} &= -\theta\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{w}_{it'''}-\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{P}_{nt'''}+\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{\kappa}_{nit'''}\right) \\
&\quad+\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{z}_{it'''}, && \text{(A.84)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{P}_{nt'''} &= \sum_i\bar{\lambda}_{nit'''}\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{w}_{it'''}+\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{\kappa}_{nit'''}-\frac{1}{\theta}\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{z}_{it'''}\right), && \text{(A.85)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{w}_{nt'''}+\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{nt'''} &= \sum_i\frac{\bar{\lambda}_{int'''}\bar{w}_{nt'''}\bar{l}_{nt'''}}{\bar{w}_{it'''}\bar{l}_{it'''}}\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{\lambda}_{int'''}+\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{w}_{it''}+\mathbb{E}_t^A\mathbb{E}_{t'}^B\mathbb{E}_{t''}^A\hat{l}_{it'''}\right). && \text{(A.86)}
\end{aligned}$$

Under Assumption 2, this system of equations simplifies to

$$\begin{aligned}
\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{v}^A_{nt'''+1} &= \frac{\partial U(\bar{w}_{nt'''+1}, \bar{P}_{nt'''+1})}{\partial \log(\bar{w}_{nt'''+1})}\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{w}_{nt'''+1} + \frac{\partial U(\bar{w}_{nt'''+1}, \bar{P}_{nt'''+1})}{\partial \log(\bar{P}_{nt'''+1})}\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{P}_{nt'''+1} \\
&\quad +\beta\sum_i \bar{\mu}_{ni,t'''}\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{v}^A_{nt'''+2}, &\text{(A.87)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\mu}^A_{ni,t'''} &= \frac{\beta}{\nu}\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{v}^A_{it'''+1} - \frac{\beta}{\nu}\sum_{m=1}^N \bar{\mu}_{nm,t'''}\beta\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{v}^A_{mt'''+1},, &\text{(A.88)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}^A_{nt'''+1} &= \sum_{i=1}^N \bar{\psi}^A_{int'''+1}\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\mu}^A_{int'''} + \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}^A_{it'''}\right), &\text{(A.89)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}^B_{nt'''+1} &= \sum_{i=1}^N \bar{\psi}^B_{int'''+1}\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\mu}^B_{int'''} + \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}^B_{it'''}\right), &\text{(A.90)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}_{nt'''+1} &= \frac{\bar{l}^A_{nt'''+1}}{\bar{l}_{nt'''+1}}\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}^A_{nt'''+1} + \frac{\bar{l}^B_{nt'''+1}}{\bar{l}_{nt'''+1}}\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}^B_{nt'''+1}, &\text{(A.91)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\lambda}_{ni,t'''} &= -\theta\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{w}_{it'} - \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{P}_{nt'''} + \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\kappa}_{nit'''}\right) + \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{z}_{it'''}, &\text{(A.92)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{P}_{nt'''} &= \sum_i \bar{\lambda}_{int'''}\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{w}_{it'} + \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\kappa}_{nit'''} - \frac{1}{\theta}\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{z}_{it'''}\right), &\text{(A.93)}\\
\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{w}_{nt'''} + \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}_{nt'''} &= \sum_i \frac{\bar{\lambda}_{int'''}\bar{w}_{nt'''}\bar{l}_{nt'''}}{\bar{w}_{it'''}\bar{l}_{it'''}}\left(\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\lambda}_{int'''} + \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{w}_{it'''} + \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}_{it'''}\right) &\text{(A.94)}
\end{aligned}$$

Given $\hat{\mu}^B_{int}$, $\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\kappa}_{nit'''}$, and $\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{z}_{it'''}$, type $A$'s decision in period $t$ is characterized by the solution to the joint system of equations (equations (A.35)-(A.42) and equations (A.87)-(A.94)) for $t' \geq t+1$ and $t''' \geq t'+1$. Following the same rationale, type $B$'s decision can be characterized by the set of equations analogous to (A.35)-(A.42) and (A.87)-(A.94).

Here we describe an algorithm to solve type $A$'s problem in period $t$.

i Given a guess of $\hat{\mu}^B_{int}$, guess expected wage and price on A's expected path, $\mathbb{E}_t^A\hat{w}_{nt'}$ and $\mathbb{E}_t^A\hat{P}_{nt'}$, and solve $A$'s decision in $t$ and expected decision in $t' > t$ using equations (A.35)-(A.36). As we are conditioning on wages and prices, in doing so we do not need $B$'s future decisions. We call the solution to this $A$'s decision in the *main branch* (solid blue line in Panel (c) of Figure 3a).

ii Given the outcome from step (i), solve the sub-branch problem described by equations (A.87)-(A.94) for all $t'$ jointly (i.e., the orange dashed lines in Panel (c) of Figure 3a) to obtain A's belief on B's decision in each $t'$, $\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\mu}^A_{int'}$, and A's belief on $B$'s belief on $A$'s migration decision in all $t'''$ (the blue dashed lines in Panel (c) of Figure 3a). Here we exploit the implication of Assumption 2 that $\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\mu}^A_{int'''} = \mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{\mu}^B_{int'''}$, namely $A$ thinks that $B$ thinks that $A$ will be convinced by the data in the next period and make the same future migration decisions.

The input to this system of equations is $\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{z}_{it'''}$, which are given outside this problem, and the belief on initial $t''' = t'$ period labor allocations $\mathbb{E}_t^A\mathbb{E}_{t'}^B\hat{l}^A_{nt'}$, which by definition is equivalent to $\mathbb{E}_t^A\hat{l}^A_{nt'}$ and is the output of Step (i).

iii Plugging $\mathbb{E}_t^A \hat{\mu}_{int'''}^A$ and $\mathbb{E}_t^A \hat{\mu}_{int'''}^B$ into the static trade equilibrium in each period along the main branch and solve for the updated $\mathbb{E}_t^A \hat{w}_{nt'}$ and $\mathbb{E}_t^A \hat{P}_{nt'}$

iv Iterate on the above steps until convergence.

The procedure explained above describes type $A$'s problem given today's $B$'s migration decision. Following the same logic, we can solve type $B$'s problem given type $A$'s migration decision today. Again, as in the proof of Proposition 4 under Assumption 1, for the solution to be an equilibrium, the belief on the other type's migration decision should be consistent with the other type's actual migration decision.

### A.4.3 Assumption 3

Let $f^g(z_{t+1}|z^t)$ be the belief of agent type $g \in \{A, B\}$ at period $t$. Under Assumption 3, agents' decisions will therefore be guided by their own (and the other types)' beliefs today, and the composite beliefs of both types in the future. Let $\tilde{f}$ be the belief (which is a probability density function) that both types think they will agree on in the future.[1] By design, in the current period, the beliefs of A and B are different. For example, given history $z^t$, type A agents' expectation on the $t+1$ period productivity is defined as before:

$$\mathbb{E}_t^A z_{t+1} \equiv \int_{z_{t+1}} z_{t+1} f^A(z_{t+1}|z^t) dz_{t+1}. \tag{A.95}$$

For $t' \geq t+2$, we use $\tilde{f}(z_{t'}|z^t)$ to weight the future outcomes up to period $t'$ and reduce the equations for future periods for each type $g \in \{A, B\}$. We denote by $\tilde{\mathbb{E}}_t z_{t'}$ the expectation on productivity where the probability weight comes from $\tilde{f}$:

$$\mathbb{E}_t^g z_{t'} = \tilde{\mathbb{E}}_t z_{t'} \equiv \int_{z_{t'}} z_{t'} \tilde{f}(z_{t'}|z^t) dz_{t'}. \tag{A.96}$$

Assumption 3 implies that higher-order beliefs of each type $g \in \{A, B\}$ and $g' \neq g$ can be simplified as

$$\tilde{\mathbb{E}}_t \hat{z}_{t''} = \mathbb{E}_t^g \hat{z}_{t''} = \mathbb{E}_t^g \mathbb{E}_{t'}^{g'} \hat{z}_{t''} \text{ for } t' \geq t+1, \text{ and } t'' \geq t'+1. \tag{A.97}$$

As in the proof under Assumption 1, we linearize around a homogeneous belief perfect foresight equilibrium and sum across the equations for all the possible states in each period weighted by $f^g(z^t)$ for $g \in \{A, B\}$ for future periods and we get the same set of equations (equations (A.35)-(A.42)). Again, equations (A.52)-(A.59) characterize the belief of type $A$ on type $B$'s future mi-

[1] One special case of this is that in each period, agents draw future types with $\rho^A$ and $\rho^B$ probability, $\rho^A + \rho^B = 1$. Let $\tilde{f}$ denote the composite density once the randomness over future draws is considered.

$$\tilde{f}(z_{t'}|z^t) \equiv \left[\rho^A f^A(z_{t'}|z^t) + \rho^B f^B(z_{t'}|z^t)\right] \text{ where } t' \geq t+2.$$

gration decision. The key difference from the proof under Assumption 1 is that for $t'' \geq t+1$, equation (A.53) is now collapsed into

$$\mathbb{E}_t^A \hat{\mu}_{ni,t''}^B = \frac{\beta}{\nu} \tilde{\mathbb{E}}_t \hat{v}_{it''+1}^B - \frac{\beta}{\nu} \sum_{m=1}^{N} \bar{\mu}_{nm,t''} \beta \tilde{\mathbb{E}}_t \hat{v}_{mt''+1}^B, \tag{A.98}$$

while the set of equations that determines the migration decision of type $B$ *in the eyes of type A agents at period t* (corresponding to equations (A.54)-(A.59)) is now given by for $t'' = t+1$,

$$\begin{aligned}
\mathbb{E}_t^A \hat{v}_{nt+1}^B &= \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{w}_{nt+1})} \mathbb{E}_t^A \hat{w}_{nt+1} + \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{P}_{nt+1})} \mathbb{E}_t^A \hat{P}_{nt+1} \\
&\quad + \beta \sum_i \bar{\mu}_{nit+1} \tilde{\mathbb{E}}_t \hat{v}_{nt+2}^B, &\text{(A.99)} \\
\mathbb{E}_t^A \hat{\mu}_{ni,t+1}^B &= \frac{\beta}{\nu} \tilde{\mathbb{E}}_t \hat{v}_{it+2}^B - \frac{\beta}{\nu} \sum_{m=1}^{N} \bar{\mu}_{nm,t''} \beta \tilde{\mathbb{E}}_t \hat{v}_{mt+2}^B, &\text{(A.100)} \\
\mathbb{E}_t^A \hat{l}_{nt+2}^A &= \sum_{i=1}^{N} \bar{\psi}_{int+2}^A \left( \mathbb{E}_t^A \hat{\mu}_{int+1}^A + \mathbb{E}_t^A \hat{l}_{it+1}^A \right), &\text{(A.101)} \\
\mathbb{E}_t^A \hat{l}_{nt+2}^B &= \sum_{i=1}^{N} \bar{\psi}_{int+2}^B \left( \mathbb{E}_t^A \hat{\mu}_{int+1}^B + \mathbb{E}_t^A \hat{l}_{it+1}^B \right), &\text{(A.102)} \\
\mathbb{E}_t^A \hat{l}_{nt+2} &= \frac{\bar{l}_{nt+2}^A}{\bar{l}_{nt+2}} \mathbb{E}_t^A \hat{l}_{nt+2}^A + \frac{\bar{l}_{nt+2}^B}{\bar{l}_{nt+2}} \mathbb{E}_t^A \hat{l}_{nt+2}^B, &\text{(A.103)} \\
\mathbb{E}_t^A \hat{\lambda}_{ni,t+1} &= -\theta \left( \mathbb{E}_t^A \hat{w}_{it'} - \mathbb{E}_t^A \hat{P}_{nt+1} + \mathbb{E}_t^A \hat{\kappa}_{ni,t+1} \right) + \mathbb{E}_t^A \hat{z}_{it+1}, &\text{(A.104)} \\
\mathbb{E}_t^A \hat{P}_{nt+1} &= \sum_i \bar{\lambda}_{ni,t+1} \left( \mathbb{E}_t^A \hat{w}_{it'} + \mathbb{E}_t^A \hat{\kappa}_{ni,t+1} - \frac{1}{\theta} \mathbb{E}_t^A \hat{z}_{it+1} \right), &\text{(A.105)} \\
\mathbb{E}_t^A \hat{w}_{nt+1} + \mathbb{E}_t^A \hat{l}_{nt+1} &= \sum_i \frac{\bar{\lambda}_{int+1} \bar{w}_{nt+1} \bar{l}_{nt+1}}{\bar{w}_{it+1} \bar{l}_{it+1}} \left( \mathbb{E}_t^A \hat{\lambda}_{int+1} + \mathbb{E}_t^A \hat{w}_{it+1} + \mathbb{E}_t^A \hat{l}_{it+1} \right), \\
& &\text{(A.106)}
\end{aligned}$$

and for $t'' \geq t+2$,

$$
\begin{aligned}
\tilde{\mathbb{E}}_t \hat{v}^B_{nt''+1} \equiv \mathbb{E}^A_t \hat{v}^B_{nt''+1} &= \frac{\partial U(\bar{w}_{nt''+1}, \bar{P}_{nt''+1})}{\partial \log(\bar{w}_{nt''+1})} \tilde{\mathbb{E}}_t \hat{w}_{nt''+1} + \frac{\partial U(\bar{w}_{nt''+1}, \bar{P}_{nt''+1})}{\partial \log(\bar{P}_{nt''+1})} \tilde{\mathbb{E}}_t \hat{P}_{nt''+1} \\
&\quad + \beta \sum_i \bar{\mu}_{nit''+1} \tilde{\mathbb{E}}_t \hat{v}^B_{nt''+2}, \qquad \text{(A.107)} \\
\tilde{\mathbb{E}}_t \hat{\mu}^B_{ni,t''} \equiv \mathbb{E}^A_t \hat{\mu}^B_{ni,t''} &= \frac{\beta}{\nu} \tilde{\mathbb{E}}_t \hat{v}^B_{it'+1} - \frac{\beta}{\nu} \sum_{m=1}^N \bar{\mu}_{nm,t''} \beta \tilde{\mathbb{E}}_t \hat{v}^B_{mt''+1}, \qquad \text{(A.108)} \\
\tilde{\mathbb{E}}_t \hat{l}^A_{nt''+1} \equiv \mathbb{E}^A_t \hat{l}^A_{nt''+1} &= \sum_{i=1}^N \bar{\psi}^A_{int''+1} \left( \tilde{\mathbb{E}}_t \hat{\mu}^A_{int''} + \tilde{\mathbb{E}}_t \hat{l}^A_{it''} \right), \qquad \text{(A.109)} \\
\tilde{\mathbb{E}}_t \hat{l}^B_{nt''+1} \equiv \mathbb{E}^A_t \hat{l}^B_{nt''+1} &= \sum_{i=1}^N \bar{\psi}^B_{int''+1} \left( \tilde{\mathbb{E}}_t \hat{\mu}^B_{int''} + \tilde{\mathbb{E}}_t \hat{l}^B_{it''} \right), \qquad \text{(A.110)} \\
\tilde{\mathbb{E}}_t \hat{l}_{nt''+1} \equiv \mathbb{E}^A_t \hat{l}_{nt''+1} &= \frac{\bar{l}^A_{nt''+1}}{\bar{l}_{nt''+1}} \tilde{\mathbb{E}}_t \hat{l}^A_{nt''+1} + \frac{\bar{l}^B_{nt''+1}}{\bar{l}_{nt''+1}} \tilde{\mathbb{E}}_t \hat{l}^B_{nt''+1}, \qquad \text{(A.111)} \\
\tilde{\mathbb{E}}_t \hat{\lambda}_{ni,t''} \equiv \mathbb{E}^A_t \hat{\lambda}_{ni,t''} &= -\theta \left( \tilde{\mathbb{E}}_t \hat{w}_{it''} - \tilde{\mathbb{E}}_t \hat{P}_{nt''} + \tilde{\mathbb{E}}_t \hat{\kappa}_{nit''} \right) + \tilde{\mathbb{E}}_t \hat{z}_{it''}, \qquad \text{(A.112)} \\
\tilde{\mathbb{E}}_t \hat{P}_{nt''} \equiv \mathbb{E}^A_t \hat{P}_{nt''} &= \sum_i \bar{\lambda}_{nit''} \left( \tilde{\mathbb{E}}_t \hat{w}_{it''} + \tilde{\mathbb{E}}_t \hat{\kappa}_{nit''} - \frac{1}{\theta} \tilde{\mathbb{E}}_t \hat{z}_{it''} \right), \qquad \text{(A.113)} \\
\mathbb{E}^A_t \hat{w}_{nt''} + \tilde{\mathbb{E}}_t \hat{l}_{nt''} \equiv \mathbb{E}^A_t \hat{w}_{nt''} + \tilde{\mathbb{E}}_t \hat{l}_{nt''} &= \sum_i \frac{\bar{\lambda}_{int''} \bar{w}_{nt''} \bar{l}_{nt''}}{\bar{w}_{it''} \bar{l}_{it''}} \left( \tilde{\mathbb{E}}_t \hat{\lambda}_{int''} + \tilde{\mathbb{E}}_t \hat{w}_{it''} + \tilde{\mathbb{E}}_t \hat{l}_{it''} \right). \qquad \text{(A.114)}
\end{aligned}
$$

For $t'' = t+1$, following the same argument as in the proof under Assumption 1, in the eye of type A agents, both types solve the same problem. For $t'' \geq t+2$, under Assumption 3, both types have the same belief for future periods, so the continuation value for both types collapses to the same set of linear equations (note that $\tilde{\mathbb{E}}_t$ is not type specific). Therefore, the future value and the migration decision for each type should coincide. As a result, we can replace $\mathbb{E}^A_t[\mu^B_{t''}$ by $\tilde{\mathbb{E}}_t[\mu^A_{t''}$ in (A.38) for $t'' > t$ as in the proof under Assumption 1. Then the system of linear equations for type $A$ can be solved analogously to Proposition 1 with type $B$'s *current* period migration decision $\hat{\mu}^B_{int}$ as given.

Type $B$'s future decisions, in turn, depends on its belief about how $A$ is going to behave in the future. Following the same argument, type $B$ believes the future value and the migration decision for type $A$ should be the same with B, which is governed by $\tilde{f}(\cdot)$. By replacing $\mathbb{E}^B_t \hat{\mu}^A_{int'}$ with $\tilde{\mathbb{E}}_t \hat{\mu}^B_{int'}$ into the equations for $B$ (analogous to equations (A.35)-(A.42)), we obtain the sequence of migration decision for type $B$ given type $A$'s *current* period migration decision $\mathbb{E}^B_t \hat{\mu}^A_{int}$.

Again, for the solution to be an equilibrium, the belief on the other type's migration decisions in $t$ should be consistent with the actual migration decision. The problem can be solved using an algorithm analogously to the one described in Section A.4.1.

### A.5 Proof of Proposition 5

#### A.5.1 Second-Order Accuracy with Heterogeneous Beliefs

**Assumption 1.** We have shown that the first-order approximation to the solution of period $t$ problem is characterized by equations (A.27)-(A.42) for $A$, taking $B$'s decision at period t as given; and analogous equations for $B$, taking A's decision in $t$ as given.

To achieve second-order accuracy in solving this problem, we can expand equations (A.27)-(A.42) to second order by incorporating type specific belief at $t$, $\mathbb{E}_t^g$ for $g \in \{A, B\}$, in equations (A.22).[2] For example, the linearized labor mobility equation and price equation in the eyes of $A$ are:

$$
\begin{aligned}
\hat{\mu}^g_{nit} &= \sum_k \frac{\beta}{\nu}\Big(\mathbb{I}(k=i) - \bar{\mu}_{nkt}\Big)\mathbb{E}_t^A \hat{v}^g_{kt+1} \\
&\quad + \frac{1}{2}\sum_{k,m}(\frac{\beta}{\nu})^2 \cdot \Big[\mathbb{I}(k=i)\bar{\mu}^g_{nit}(\mathbb{I}(m=i) - \bar{\mu}^g_{nmt}) \\
&\quad - \bar{\mu}^g_{nkt}\bar{\mu}^g_{nit}[\mathbb{I}(m=i) - 2\bar{\mu}^g_{nmt} + \mathbb{I}(m=k)]\Big]\mathbb{E}_t^A \hat{v}^g_{kt+1}\hat{v}^g_{mt+1} \\
\mathbb{E}_t^A \hat{P}_{nt} &= \sum_{i=1}^N \bar{\lambda}_{nit}\mathbb{E}_t^A \hat{w}_{it} + \sum_{i=1}^N \bar{\lambda}_{nit} \cdot \mathbb{E}_t^A \hat{w}_{it} + \sum_{i=1}^N \Big(-\frac{1}{\theta}\Big)\bar{\lambda}_{nit}\mathbb{E}_t^A \hat{z}_{it} \\
&\quad + \frac{1}{2}\sum_{i,m}\theta\Big(-\mathbf{I}(i=m)\bar{\lambda}_{nit} + \bar{\lambda}_{nit}\bar{\lambda}_{nmt}\Big) \cdot \Big(\mathbb{E}_t^A \hat{w}_{it}\hat{w}_{mt} + \mathbb{E}_t^A \hat{\kappa}_{nit}\hat{\kappa}_{nmt} + \Big(\frac{1}{\theta}\Big)^2 \mathbb{E}_t^A \hat{z}_{it}\hat{z}_{mt}\Big) \\
&\quad + \sum_{i,m}\theta\Big(-\mathbf{I}(i=m)\bar{\lambda}_{nit} + \bar{\lambda}_{nit}\bar{\lambda}_{nmt}\Big)\Big(\mathbb{E}_t^A \hat{w}_{it}\hat{\kappa}_{nmt} - \Big(\frac{1}{\theta}\Big)\mathbb{E}_t^A \hat{w}_{it}\hat{z}_{mt} - \Big(\frac{1}{\theta}\Big)\mathbb{E}_t^A \hat{\kappa}_{nit}\hat{z}_{mt}\Big)
\end{aligned}
\tag{A.115}
$$

The full set of equations for type $A$'s problem are given in (A.116)-(A.117).

[2] Here we only display $A$'s problem but $B$'s problem can be described analogously.

For each $g \in \{A, B\}$,

$$\begin{aligned}
\mathbb{E}_t^A \hat{v}^g_{nt+1} &= \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{w}_{nt+1})}\mathbb{E}_t^A \hat{w}_{nt+1} + \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{P}_{nt+1})}\mathbb{E}_t^A \hat{P}_{nt+1} \\
&+ \frac{\partial^2 u(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{w}_{nt+1})^2}\mathbb{E}_t^A \hat{w}^2_{nt+1} + \frac{\partial^2 u(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{P}_{nt+1})^2}\mathbb{E}_t^A \hat{P}^2_{nt+1} \\
&+ \sum_k \beta \bar{\mu}^g_{nkt+1}\mathbb{E}_t^A \hat{v}^g_{kt+2} + \frac{1}{2}\sum_m \sum_k \frac{\beta^2}{\nu}\bar{\mu}^g_{nmt+1}\Big(\mathbb{I}(k=m) - \bar{\mu}^g_{nkt+1}\Big)\mathbb{E}_t^A \hat{v}^g_{mt+2}\hat{v}^g_{kt+2} \\
\hat{\mu}^g_{nit} &= \sum_k \frac{\beta}{\nu}\Big(\mathbb{I}(k=i) - \bar{\mu}_{nkt}\Big)\mathbb{E}_t^A \hat{v}^g_{kt+1} \\
&+ \frac{1}{2}\sum_{k,m}(\frac{\beta}{\nu})^2 \cdot \Big[\mathbb{I}(k=i)\bar{\mu}^g_{nit}(\mathbb{I}(m=i) - \bar{\mu}^g_{nmt}) - \bar{\mu}^g_{nkt}\bar{\mu}^g_{nit}[\mathbb{I}(m=i) - 2\bar{\mu}^g_{nmt} + \mathbb{I}(m=k)]\Big]\mathbb{E}_t^A \hat{v}^g_{kt+1}\hat{v}^g_{mt+1} \\
\hat{l}^g_{nt+1} &= \sum_{k=1}^N \frac{\beta}{\nu}\sum_{i=1}^N \bar{\psi}^g_{int+1}\Big(\mathbb{I}(k=n) - \bar{\mu}^g_{ikt}\Big)\mathbb{E}_t^A \hat{v}^A_{kt+1} + \sum_{k=1}^N \bar{\psi}^g_{knt+1}\mathbb{E}_t^A \hat{l}^g_{kt} \\
&+ \frac{1}{2}\sum_{k,m}(\frac{\beta}{\nu})^2 \cdot \Big[ - \sum_{i=1}^N \bar{\mu}^g_{ikt}\bar{\psi}^g_{int+1}[\mathbb{I}(m=n) + \mathbb{I}(m=k) - 2\bar{\mu}^g_{imt} + \mathbb{I}(k=n)\sum_i \bar{\psi}^g_{int+1}(\mathbb{I}(m=n) - \bar{\mu}^g_{imt}) \\
&- [\sum_{i=1}^N \bar{\psi}^g_{int+1}(\mathbb{I}(k=n) - \bar{\mu}^g_{ikt})] \cdot [\sum_i \bar{\psi}^g_{int+1}(\mathbb{I}(m=n) - \bar{\mu}^g_{imt})]\Big]\mathbb{E}_t^A \hat{v}^g_{kt+1}\hat{v}^g_{mt+1} \\
&+ \frac{1}{2}\sum_{k,m}\Big[\mathbb{1}(k=m)\bar{\psi}^g_{mnt+1} - \bar{\psi}^g_{knt+1}\bar{\psi}^g_{mnt+1}\Big]\mathbb{E}_t^A \hat{l}^g_{kt}\hat{l}^g_{mt} \\
&+ \sum_{k,m}\Big[\frac{\beta}{\nu}\bar{\psi}^g_{mnt+1}(\mathbb{I}(k=n) - \bar{\mu}^g_{mkt}) - \frac{\beta}{\nu}\bar{\psi}^g_{mnt+1}[\sum_i \bar{\psi}^g_{int+1}(\mathbb{I}(k=n) - \bar{\mu}^g_{ikt})]\Big]\mathbb{E}_t^A \hat{v}^g_{kt+1}\hat{l}^g_{mt} \\
\hat{l}_{nt+1} &= \frac{\bar{l}^A_{nt+1}}{\bar{l}_{nt+1}}\hat{l}^A_{nt+1} + \frac{\bar{l}^B_{nt+1}}{\bar{l}_{nt+1}}\hat{l}^B_{nt+1}, \\
\mathbb{E}_t^A \hat{\lambda}_{nit} &= -\theta(\mathbb{E}_t^A \hat{w}_{it} + \mathbb{E}_t^A \hat{\kappa}_{int} - \mathbb{E}_t^A \hat{P}_{nt}) + \mathbb{E}_t^A \hat{z}_{it} \\
\mathbb{E}_t^A \hat{P}_{nt} &= \sum_{i=1}^N \bar{\lambda}_{nit}\mathbb{E}_t^A \hat{w}_{it} + \sum_{i=1}^N \bar{\lambda}_{nit} \cdot \mathbb{E}_t^A \hat{w}_{it} + \sum_{i=1}^N \Big(-\frac{1}{\theta}\Big)\bar{\lambda}_{nit}\mathbb{E}_t^A \hat{z}_{it} \\
&+ \frac{1}{2}\sum_{i,m}\theta\Big(-\mathbf{I}(i=m)\bar{\lambda}_{nit} + \bar{\lambda}_{nit}\bar{\lambda}_{nmt}\Big) \cdot \Big(\mathbb{E}_t^A \hat{w}_{it}\hat{w}_{mt} + \mathbb{E}_t^A \hat{\kappa}_{nit}\hat{\kappa}_{nmt} + \Big(\frac{1}{\theta}\Big)^2 \mathbb{E}_t^A \hat{z}_{it}\hat{z}_{mt}\Big) \\
&+ \sum_{i,m}\theta\Big(-\mathbf{I}(i=m)\bar{\lambda}_{nit} + \bar{\lambda}_{nit}\bar{\lambda}_{nmt}\Big)\Big(\mathbb{E}_t^A \hat{w}_{it}\hat{\kappa}_{nmt} - \Big(\frac{1}{\theta}\Big)\mathbb{E}_t^A \hat{w}_{it}\hat{z}_{mt} - \Big(\frac{1}{\theta}\Big)\mathbb{E}_t^A \hat{\kappa}_{nit}\hat{z}_{mt}\Big) \\
\mathbb{E}_t^A \hat{w}_{nt} + \mathbb{E}_t^A \hat{l}_{nt} &= \sum_{i=1}^N \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{\bar{w}_{nt}\bar{l}_{nt}}\Big(\mathbb{E}_t^A \hat{\lambda}_{int} + \mathbb{E}_t^A \hat{w}_{it} + \mathbb{E}_t^A \hat{l}_{it}\Big) \\
&+ \frac{1}{2}\sum_{i,m}\Big(1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})} - \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}\Big)\Big(\mathbb{E}_t^A \hat{w}_{it}\hat{w}_{mt} + \mathbb{E}_t^A \hat{\lambda}_{int}\hat{\lambda}_{mnt} + \mathbb{E}_t^A \hat{l}_{it}\hat{l}_{mt}\Big) \\
&+ \sum_{i,m}\Big(1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})} - \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}\Big)\Big(\mathbb{E}_t^A \hat{\lambda}_{int}\hat{w}_{mt} + \mathbb{E}_t^A \hat{\lambda}_{int}\hat{l}_{mt} + \mathbb{E}_t^A \hat{w}_{it}\hat{l}_{mt}\Big). \qquad \text{(A.116)}
\end{aligned}$$

For $t' \geq t+1$,

$$\begin{aligned}
\mathbb{E}_t^A \hat{v}^A_{nt'+1} &= \frac{\partial U(\bar{w}_{nt'+1}, \bar{P}_{nt'+1})}{\partial \log(\bar{w}_{nt'+1})} \mathbb{E}_t^A \hat{w}_{nt'+1} + \frac{\partial U(\bar{w}_{nt'+1}, \bar{P}_{nt'+1})}{\partial \log(\bar{P}_{nt'+1})} \mathbb{E}_t^A \hat{P}_{nt'+1} \\
&+ \frac{\partial^2 u(\bar{w}_{nt'+1}, \bar{P}_{nt'+1})}{\partial \log(\bar{w}_{nt'+1})^2} \mathbb{E}_t^A \hat{w}^2_{nt'+1} + \frac{\partial^2 u(\bar{w}_{nt'+1}, \bar{P}_{nt'+1})}{\partial \log(\bar{P}_{nt'+1})^2} \mathbb{E}_t^A \hat{P}^2_{nt'+1} \\
&+ \sum_k \beta \bar{\mu}^A_{nkt'+1} \mathbb{E}_t^A \hat{v}^A_{kt+2} + \frac{1}{2} \sum_m \sum_k \frac{\beta^2}{\nu} \bar{\mu}^A_{nmt'+1} \Big( \mathbb{I}(k=m) - \bar{\mu}^A_{nkt'+1} \Big) \mathbb{E}_t^A \hat{v}^A_{mt+2} \hat{v}^A_{kt+2} \\
\mathbb{E}_t^A \hat{\mu}^g_{nit'} &= \sum_k \frac{\beta}{\nu} \Big( \mathbb{I}(k=i) - \bar{\mu}^g_{nkt} \Big) \mathbb{E}_t^A \hat{v}^g_{kt'+1} \\
&+ \frac{1}{2} \sum_{k,m} (\frac{\beta}{\nu})^2 \Big[ \mathbb{I}(k=i) \bar{\mu}^g_{nit'} (\mathbb{I}(m=i) - \bar{\mu}^g_{nmt'}) \\
&\qquad - \bar{\mu}^g_{nkt} \bar{\mu}^g_{nit'} [\mathbb{I}(m=i) - 2\bar{\mu}^g_{nmt'} + \mathbb{I}(m=k)] \Big] \mathbb{E}_t^A \hat{v}^g_{kt'+1} \hat{v}^g_{mt'+1} \\
\mathbb{E}_t^A \hat{l}^g_{nt'+1} &= \sum_{k=1}^N \frac{\beta}{\nu} \sum_{i=1}^N \bar{\psi}^g_{int'+1} \Big( \mathbb{I}(k=n) - \bar{\mu}^g_{ikt} \Big) \mathbb{E}_t^A \hat{v}^g_{kt'+1} + \sum_{k=1}^N \bar{\psi}^g_{knt'+1} \mathbb{E}_t^A \hat{l}^g_{kt} \\
&+ \frac{1}{2} \sum_{k,m} (\frac{\beta}{\nu})^2 \Big[ - \sum_{i=1}^N \bar{\mu}^g_{ikt} \bar{\psi}^g_{int'+1} [\mathbb{I}(m=n) + \mathbb{I}(m=k) - 2\bar{\mu}^g_{imt'} + \mathbb{I}(k=n) \sum_i \bar{\psi}^g_{int'+1} (\mathbb{I}(m=n) - \bar{\mu}^g_{imt'}) \\
&- \big[ \sum_{i=1}^N \bar{\psi}^g_{int'+1} \big( \mathbb{I}(k=n) - \bar{\mu}^g_{ikt} \big) \big] \big[ \sum_i \bar{\psi}^g_{int'+1} \big( \mathbb{I}(m=n) - \bar{\mu}^g_{imt'} \big) \big] \Big] \mathbb{E}_t^A \hat{v}^g_{kt'+1} \hat{v}^g_{mt'+1} \\
&+ \frac{1}{2} \sum_{k,m} \mathbf{1}(k=m) \bar{\psi}^g_{mnt'+1} - \bar{\psi}^g_{knt'+1} \bar{\psi}^g_{mnt'+1} \mathbb{E}_t^A \hat{l}^g_{kt'} \hat{l}^g_{mt'} \\
&+ \sum_{k,m} \frac{\beta}{\nu} \bar{\psi}^A_{mnt'+1} \big( \mathbb{I}(k=n) - \bar{\mu}^A_{mkt} \big) - \frac{\beta}{\nu} \bar{\psi}^A_{mnt'+1} \big[ \sum_i \bar{\psi}^A_{int'+1} \big( \mathbb{I}(k=n) - \bar{\mu}^A_{ikt} \big) \big] \mathbb{E}_t^A \hat{v}^A_{kt'+1} \hat{l}^A_{mt'}, \\
\mathbb{E}_t^A \hat{l}_{nt'+1} &= \frac{\bar{l}^A_{nt'+1}}{\bar{l}_{nt'+1}} \mathbb{E}_t^A \hat{l}^A_{nt'+1} + \frac{\bar{l}^B_{nt'+1}}{\bar{l}_{nt'+1}} \mathbb{E}_t^A \hat{l}^B_{nt'+1}, \\
\mathbb{E}_t^A \hat{\lambda}_{nit'} &= -\theta (\mathbb{E}_t^A \hat{w}_{it'} + \mathbb{E}_t^A \hat{\kappa}_{int'} - \mathbb{E}_t^A \hat{P}_{nt'}) + \mathbb{E}_t^A \hat{z}_{it'} \\
\mathbb{E}_t^A \hat{P}_{nt'} &= \sum_{i=1}^N \bar{\lambda}_{nit'} \mathbb{E}_t^A \hat{w}_{it'} + \sum_{i=1}^N \bar{\lambda}_{nit'} \mathbb{E}_t^A \hat{w}_{it'} + \sum_{i=1}^N \Big( -\frac{1}{\theta} \Big) \bar{\lambda}_{nit'} \mathbb{E}_t^A \hat{z}_{it'} \\
&+ \frac{1}{2} \sum_{i,m} \theta \Big( -\mathbf{I}(i=m) \bar{\lambda}_{nit'} + \bar{\lambda}_{nit'} \bar{\lambda}_{nmt'} \Big) \Big( \mathbb{E}_t^A \hat{w}_{it'} \hat{w}_{mt'} + \mathbb{E}_t^A \hat{\kappa}_{nit'} \hat{\kappa}_{nmt'} + \Big( \frac{1}{\theta} \Big)^2 \mathbb{E}_t^A \hat{z}_{it'} \hat{z}_{mt'} \Big) \\
&+ \sum_{i,m} \theta \Big( -\mathbf{I}(i=m) \bar{\lambda}_{nit'} + \bar{\lambda}_{nit'} \bar{\lambda}_{nmt'} \Big) \Big( \mathbb{E}_t^A \hat{w}_{it'} \hat{\kappa}_{nmt'} - \Big( \frac{1}{\theta} \Big) \mathbb{E}_t^A \hat{w}_{it'} \hat{z}_{mt'} - \Big( \frac{1}{\theta} \Big) \mathbb{E}_t^A \hat{\kappa}_{nit'} \hat{z}_{mt'} \Big) \\
\mathbb{E}_t^A \hat{w}_{nt'} + \mathbb{E}_t^A \hat{l}_{nt'} &= \sum_{i=1}^N \frac{\bar{\lambda}_{int'} \bar{w}_{it'} \bar{l}_{it'}}{\bar{w}_{nt'} \bar{l}_{nt'}} \Big( \mathbb{E}_t^A \hat{\lambda}_{int'} + \mathbb{E}_t^A \hat{w}_{it'} + \mathbb{E}_t^A \hat{l}_{it'} \Big) \\
&+ \frac{1}{2} \sum_{i,m} \Big( 1(m=i) \frac{\bar{\lambda}_{int'} \bar{w}_{it'} \bar{l}_{it'}}{(\bar{w}_{nt'} \bar{l}_{nt'})} - \frac{\bar{\lambda}_{int'} \bar{w}_{it'} \bar{l}_{it'} \bar{\lambda}_{mnt'} \bar{w}_{mt'} \bar{l}_{mt'}}{(\bar{w}_{nt'} \bar{l}_{nt'})^2} \Big) \Big( \mathbb{E}_t^A \hat{w}_{it'} \hat{w}_{mt'} + \mathbb{E}_t^A \hat{\lambda}_{int'} \hat{\lambda}_{mnt'} + \mathbb{E}_t^A \hat{l}_{it'} \hat{l}_{mt'} \Big) \\
&+ \sum_{i,m} \Big( 1(m=i) \frac{\bar{\lambda}_{int'} \bar{w}_{it'} \bar{l}_{it'}}{(\bar{w}_{nt'} \bar{l}_{nt'})} - \frac{\bar{\lambda}_{int'} \bar{w}_{it'} \bar{l}_{it'} \bar{\lambda}_{mnt'} \bar{w}_{mt'} \bar{l}_{mt'}}{(\bar{w}_{nt'} \bar{l}_{nt'})^2} \Big) \Big( \mathbb{E}_t^A \hat{\lambda}_{int'} \hat{w}_{mt'} + \mathbb{E}_t^A \hat{\lambda}_{int'} \hat{l}_{mt'} + \mathbb{E}_t^A \hat{w}_{it'} \hat{l}_{mt} \Big)
\end{aligned} \tag{A.117}$$

where $\bar{\psi}^g_{int+1} = \frac{\bar{\mu}^g_{int} \bar{l}^g_{it}}{\bar{l}^g_{nt+1}}$.

We propose the following algorithm to show that this problem is tractable.

1. For each period $t$, simulate a set of $S$ paths of fundamentals from $t$ to $T$ using each type's belief, $f^A(z^T|z^t) \equiv f^A(z^T|z^{T-1}) \cdot f^A(z^{T-1}|z^{T-2})...f^A(z^{t+1}|z^t)$ and $f^B(z^T|z^t) \equiv f^B(z^T|z^{T-1}) \cdot f^B(z^{T-1}|z^{T-2})...f^B(z^{t+1}|z^t)$, respectively. Given the other type's current period migration decisions, each type's problem is analogous to equations (A.22), and the solution algorithm used in the first-order equations can be identically applied. Here, the initial fundamental at period $t$ is identical across the S paths. Denote by $z_{t'}^{tsA}$ the $s^{th}$ simulated productivity for period $t'$ generated at period $t$ using $A$'s belief.

2. We solve the current period decision of each type, $\mu_{int}^A$ and $\mu_{int}^B$, respectively, up to first-order. To do so, generate the mean belief path on fundamentals under each belief, $\mathbb{E}_t^g z^{t'}$ for $g \in \{A, B\}$ and all $t' \geq t$, and solve for the solution under the mean path. Note that note that across the S paths, there is only one decision for A and one decision for B for period $t$. Simulations matter only for outcomes in $t' > t$.

3. For each type, solve the problem sequentially for $t$ to $T$ up to first-order for each simulation. One feasible algorithm to solve this is:

   (a) Compute the belief on fundamental for each simulation $s$ as a deviation from the mean belief path on fundamental, i.e., $\mathbb{E}_t^A \hat{z}_{t'}^{tsA} \equiv \log z_{t'}^{tsA} - \log \mathbb{E}_t^A z^{t'}$ for type $A$.

   (b) Solve for the solution under each simulation as deviations from the mean path - this can be done using the proof of Proposition 4 under Assumption 1 (Section A.4.1). Here, due to Assumption 1, when solving future decisions, we essentially assume homogeneous belief model.

   We obtain the expected second-order moments, e.g., $\mathbf{E}_t^A v_{nt'} v_{kt'}$, by taking the average of the solutions for $S$ simulated paths for A and B separately.

4. Plug the second-order moments obtained from step 3 into equations (A.116)-(A.117). Then solve the system of equations for type $A$ and type $B$ characterizing Nash Equilibrium again. The equilibrium obtained this way should have second-order accuracy.

**Remark.** The above algorithm means that we don't have to simulate $S$ equilibria in the decision of agents in period $t$ as described in Step (iv) of the algorithm in Section A.4.1. Instead, we only need to solve $2 \times S$ homogeneous beliefs problems, and two equilibriums (one for obtaining the decision today before doing the simulation for the future, one at last to ensure the solution to the equilibrium has second-order accuracy).

**Assumption 2.** This is computationally demanding, but conceptually it follows from the first-order algorithm in the same way.

1. First, solve for the migration decisions of $A$ and $B$ for current period $t$, with first-order accuracy.

2. Simulate future *tree* (not just paths). That is, simulate a set of paths of fundamentals with the main branch from $f^A(z^T|z^t)$ (solid lines in Figure 3a), and the off-branch limbs from

$f^B(z^T|z^{t'}) \cdot f^A(z^{t'}|z^t)$ for all $t' \in \{t+1, ..., T\}$ for $A$ (dashed lines in Figure 3a) and analogously for $B$. Note now the total number of simulations required grow quadratically with $T$.

3. For each simulated main branch for each type, solve the corresponding off-branch outcomes for many sub-simulations at a time up to first-order. One feasible algorithm to solve each simulated *tree* is proposed in the proof of Proposition 4 under Assumption 2. This ensures internal consistency in the main branch and the 'limbs' for a solution at each of the sub branches. Calculate the expected second-order moments based on these simulations.

4. Plug in the second-order moments for $A'$ and $B$ to their own system of equations, and solve the joint system of equations for the remaining terms. The solution should have second-order accuracy.

**Assumption 3.** The procedure proceeds in a similar way to under Assumption 1—in fact it is a special case of Assumption 1.

1. Simulate a set of S paths of fundamentals from $t$ to $T$ using the 'tilde' belief, $\tilde{f}(z^T|z_t)$, defined in A.96. Under Assumption 3, the belief of A and B at period $t$ about $t' > t+1$ are the same.

2. We solve the Nash equilibrium to obtain the decision of the current period for each type, $\mu^A_{int}$ and $\mu^B_{int}$, respectively, up to first-order. Under Assumption 3, each type's belief about the next period is different.

3. For each type, solve the problem sequentially for t to T up to first-order for each simulation using the algorithm proposed in proof of Proposition 4 under Assumption 3 (Section A.4.3). Calculate the expected second-order moments, e.g., $\mathbf{E}^A_t v^A_{nt'} v^A_{kt'}$, by taking the average of the solutions for $S$ simulated paths for A and B separately.

4. Plug A's second order terms and B's second order terms to their own equations, (A.116)-(A.117) for $A$ and analogous for $B$. Then solve the system of equations for type $A$ and type $B$ characterizing Nash Equilibrium again.

### A.5.2 Second-Order Accuracy in Ex-Post Evaluations

Without loss of generality, assume researchers observe the realizations over $t = 1, 2, ..., T$, denoted by $\bar{x}_{it}$, where $x_{it} \in \{v_{it}, P_{it}, w_{it}, \mu_{int}, \lambda_{int}, l_{nt}\}$. We would like to know if agents' have perfect foresight, what would the realization be like with second-order accuracy.[3] To achieve this goal, we should acknowledge that agents face uncertainty in making decisions and that their choice reflect such uncertainty.

The algorithm described in Section 2.5 can be readily generalized to accommodate second-order approximation. Specifically, in Step (i) and (iii) of the algorithm, instead of solving for the

[3] We might also be interested in counterfactuals. But once we have the PF path corresponding to that particular ex-post observation, we can implement counterfactuals using Proposition 3.

expected outcomes with first-order accuracy using Proposition 1, use Proposition 3 to solve for the expected outcomes with second-order accuracy.[4]

From these expected outcomes starting from each period $t = 1, ..., T-1$, we can solve for the PF path as deviations from these expected paths with second-order accuracy by again using Proposition 3 with a twist—instead of taking the PF path as given, now the PF path is the output, and the expected values (with second-order accuracy) and covariances from simulations are inputs.

### A.5.3 Heterogeneous Beliefs in Ex-Post Evaluations

Suppose researchers observe the evolution of outcomes from $t = 1$ to $T$. Researchers impose that the economy features heterogeneous beliefs and are interested in counterfactual questions such as how the economy would evolve if agents held different beliefs from the ones that generate the data, or if the fundamental productivity processes have a different realization. We will consider two polar cases in terms of what data researchers have access to.

**Strong data requirement.** We start with what we call 'strong data requirement,' strong in the sense that researchers observe all past decisions by belief type. To rationalize such data, we need to assume that the two types of agents differ in their (potentially time-varying) migration costs. Intuitively, while the model can generate heterogeneous migration through belief alone, if we were to use any external measure to discipline beliefs, there will not be enough degree of freedom to rationalize migration unless we allow migration costs to differ.

We extend Proposition 4 to accommodate different migration costs between the types, approximating around a deterministic sequential path with two types of agents that differ in their of migration costs. The approximation to the market clearing condition is the same as in Proposition 4. The approximation to migration decisions is type specific: i.e., A's decision (and B's belief about A's decision and so on) is approximated around A's decision in the PF sequential path; B's decision (and A's belief about B's decision and so on) is approximated around B's decision in the PF sequential path. This way, we do not have to back out the migration costs as all of them will be canceled within a type-period.

**Ex-post algorithm.** Starting from the sequence of data, we develop a recursive algorithm to recover both types of agents' beliefs about future prices and quantities, which extends the algorithm described in Section 2.5:

i Solve for the expected outcomes in period $T$ according to agents' belief in T-1, i.e., $\mathbb{E}^g_{T-1}(w_{nT})$, $\mathbb{E}^g_{T-1}(P_{nT})$, $\mathbb{E}^g_{T-1}(\lambda_{nT})$ for $g \in \{A, B\}$ using (A.32)-(A.34) for $t = T-1$ (and analogous equations for $B$).

ii Appending outcome from step (i) with A and B's actual migration in period $T-1$, $\mu^A_{inT-1}$ and $\mu^B_{inT-1}$, we then obtain the solution to the joint system of equations at period $T-1$ defined in Section A.4. Note that as $T-1$ is the second-to-the-last period, higher-order beliefs do

[4] This involves simulating deviations and solving for the outcomes under these simulated deviations. Then calculate expected second-order moments, and use these to calculate the expected outcomes with second-order accuracy.

not matter for decisions in $T-1$. So solving the problem in period $T-1$ around a PF path is essentially solving equations in Proposition 1 extended to incorporate two types of agents.

iii Solve for the $\mathbb{E}^g_{T-2}(w_{nT-1})$, $\mathbb{E}^g_{T-2}(w_{nT})$, $\mathbb{E}^g_{T-2}(P_{nT-1})$, $\mathbb{E}^g_{T-2}(P_{nT})$, $\mathbb{E}^g_{T-2}(\mu^A_{nT-1})$ for $g \in \{A, B\}$ by approximating around the outcome from Step (ii) in the previous step. This is equivalent to solving (A.27)-(A.42) for $t = T-2$ with the approximation points (bar values in equations (A.27)-(A.42)) coming from $\mathbb{E}^g_{T-1}(w_{nT})$, $\mathbb{E}^g_{T-1}(P_{nT})$, $\mu^g_{inT-1}$, $w_{nT-1}$, $P_{nT-1}$, and $l_{nT-1}$, for $g \in \{A, B\}$. The input will be the difference in first- and second-order beliefs between T-2 and T-1 (for A and B, respectively). Solving this under each assumption on high-order belief can be done analogously to the process explained in the proof of Proposition 4. The output of this step will be in deviations. These deviations, when multiplied by the level from the previous step, give us the expected value of outcomes according to beliefs at $T-2$.

iv Append the above with the actual migration decisions of $A$ and $B$ in period $T-2$, this gives us the solution to the problem of A's and B's in their expected values.

v Iterate on Steps (iii) and (iv) from $T-3$ backward until $t = 1$

**Weak Data Requirement**. As a polar case, assume that researchers only observe the total migration (the sum across belief types) between each pair of locations. In addition, researchers have access to one cross-sectional information on the share of believers at each location in period $T$ (we also discuss if researchers only have this information in other periods). We assume that agents have the same migration costs. In fact, as shown below, our model is just identified if agents have the same migration costs and researchers have only one cross-sectional of the data (no matter which period). If researchers have more than one cross-sectional distribution of the population by belief, the model will be over-identified.

i Solve for $\mathbb{E}^g_{T-1}(w_{nT})$, $\mathbb{E}^g_{T-1}(P_{nT})$, $\mathbb{E}^g_{T-1}(\lambda_{nT})$ for $g \in \{A, B\}$ using (A.32)-(A.34) for $t = T-1$ (and analogous equations for $B$).

ii Use output from Step (i) to derive the difference in the migration of A and B. The intuition is that the difference in A and B's migration decision in period $T-1$ is entirely their belief about $T$ outcomes, which we have derived in the previous step. Thus, the difference between $\mathbb{E}^A_{T-1}U(w_{nT}/p_{nT})$, and $\mathbb{E}^B_{T-1}U(w_{nT}/p_{nT})$ gives us information on the difference in the migration behavior between A and B.

To implement this idea, linearize A's decision around B's decision (which is an unobserved $N \times N$ matrix, denoted by $\mu^B_{T-1}$ for now). A's decision will be a linear function of $\mu^B_{T-1}$ and the difference between $\mathbb{E}^A_{T-1}U(w_{nT}/p_{nT})$, and $\mathbb{E}^B_{T-1}U(w_{nT}/p_{nT})$. Note also that the weighted sum of $\mu^B_{T-1}$ and $\mu^A_{T-1}$ should jointly be equal to the total $N \times N$ migration matrix in the data. Let variables with a bar indicate the data, which the researcher observed.

$$\hat{\mu}^A_{nm,T-1} = \frac{\beta}{\nu}\sum_{k=1}^{N}\left(\mathbb{1}(k=m) - \mu^B_{nkT-1}\right)\cdot[\mathbb{E}^A_{T-1}U(w_{kT}/p_{kT}) - \mathbb{E}^B_{T-1}U(w_{kT}/p_{kT})] \tag{A.118}$$

$$\begin{aligned}
\mu^A_{nm,T-1} &\equiv \exp(\hat{\mu}^A_{nm,T-1})\cdot\mu^B_{nmT-1} \\
\bar{\mu}_{nm,T-1}\bar{l}_{n,T-1} &= \mu^A_{nm,T-1}l^A_{nT-1} + \mu^B_{nm,T-1}l^B_{n,T-1} \\
\bar{l}^A_{nT} &= \sum_i^N \mu^A_{inT-1}\cdot l^A_{iT-1} \\
\bar{l}^B_{nT} &= \sum_i^N \mu^B_{inT-1}\cdot l^B_{iT-1}
\end{aligned}$$

where the right-hand side of the first equation is due to $\hat{v}$ being level-difference and $T$ being the last period, and the third equation simply says that total migration is consistent with the decomposition by belief type.

This system of equations has $N\times N\times 2 + N\times 2$ unknown (with normalization), and the same number of equations, so it has a unique solution. This shows that given the belief structure, there is a unique way of rationalizing the observed migration patterns and the distribution of agents by belief in $T$.

This solution also gives us the distribution of agents by belief in period $T-1$. We can use these to recover agents' distribution recursively backward as below.

iii Recover the belief of agents in period $T-2$ about outcomes in period $T-1$ and forward, following Step (iii) of the algorithm for heterogeneous belief case described at the beginning of this subsection (**'Ex-post algorithm'**).

iv From Step (iii), we obtain $\mathbb{E}^g_{T-2}v^{g'}_{nT-1}$ for $g, g' \in \{A, B\}$. Use the value difference between A's and B's belief in Step (iii) to derive the migration decisions analogously to equations (A.118). For example, A's migration decision for $T-2$ should satisfy:

$$\hat{\bar{\mu}}^A_{nm,T-2} = \frac{\beta}{\nu}\sum_{k=1}^{N}\left(\mathbb{1}(k=m) - \bar{\mu}^B_{nkT-2}\right)\cdot[\mathbb{E}^A_{T-2}v^A_{T-1} - \mathbb{E}^B_{T-2}v^B_{T-1}].$$

Append the outcome from Step (iii) with the outcome from Step (iv).

v Iterate on Steps (iii) and (iv) until t=1.

In the above algorithm, we assume that the researchers observed the distribution of the population by their belief type in $T$. If instead the researchers observed the distribution in $t' \neq T$, we can nest the above algorithm within an outer loop, in which we guess a distribution by belief type in period $T$, iterate on the above algorithm until the guess implies consistent distribution in period $t'$ as the data.

## A.6 Additional Derivations

In this subsection, we derive the first and second-order approximations of equations (2)-(7).

### A.6.1 First-Order Derivatives

**First-order Approximation to value function**

Let $v_{nt+1} = \left(U(w_{nt+1}, P_{nt+1}) + \nu \log\left[\sum_{i=1}^{N} \exp(\beta v_{it+2}(z^{t+1}) - m_{nit+1})^{1/\nu}\right]\right) \equiv F_n(\omega_{it+1}, \mathbf{v_{t+2}})$. Then, a first-order approximation of $\mathbb{E}_t\hat{v}_{nt+1}$ is:

$$\begin{aligned}
\mathbb{E}_t\hat{v}_{nt+1} \quad &\approx \quad \frac{\partial F_n}{\partial \log(\bar{w}_{nt+1})}\mathbb{E}_t\hat{w}_{nt+1} + \frac{\partial F_n}{\partial \log(\bar{P}_{nt+1})}\mathbb{E}_t\hat{P}_{nt+1} + \frac{\partial F_n}{\partial \mathbf{\bar{v}_{t+2}}}\mathbb{E}_t\mathbf{\hat{v}_{t+2}}(z^{t+1}) \\
&= \quad \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{w}_{nt+1})}\mathbb{E}_t\hat{w}_{nt+1}(z^{t+1}) + \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{P}_{nt+1})}\mathbb{E}_t\hat{P}_{nt+1}(z^{t+1}) + \beta\sum_{i=1}^{N}\bar{\mu}_{nit+1}\mathbb{E}_t\hat{v}_{it+2}(z^{t+1})
\end{aligned} \tag{A.119}$$

**First-order Approximation to mobility flows**

Let $\log \mu_{nit}(z^t) = \log\left(\frac{\exp\left(\beta v_{it+1}(z^t) - m_{nit}\right)^{1/\nu}}{\sum_{h=1}^{N}\exp(\beta v_{ht+1}(z^t) - m_{nht})^{1/\nu}}\right) \equiv G_{ni}(\mathbf{v_{t+1}}(z^t))$. Then, a first-order approximation of $\mathbb{E}_t\hat{\mu}_{nit}$ is:

$$\begin{aligned}
\mathbb{E}_t\hat{\mu}_{nit} \quad &\approx \quad \frac{\partial G_{ni}}{\partial \mathbf{\bar{v}_{t+1}}}\mathbb{E}_t\mathbf{\hat{v}_{t+1}} \\
&= \quad \frac{\beta}{\nu}\sum_{k=1}^{N}\Big(\mathbb{1}(k=i) - \bar{\mu}_{nkt}\Big)\mathbb{E}_t\hat{v}_{kt+1}
\end{aligned} \tag{A.120}$$

**First-order Approximation to law of motion of labor**

Let $\log l_{nt+1} = \log\sum_{i}^{N}\mu_{int}l_{it}(z^{t-1}) \equiv J(\boldsymbol{\mu_{nt+1}}, \mathbf{l_t}(z^{t-1}))$ where $\boldsymbol{\mu_{nt}} = [\mu_{1nt}, ..., \mu_{Nnt}]'$. Then, a first-order approximation of $\mathbb{E}_t\hat{l}_{nt+1}$ is:

$$\begin{aligned}
\mathbb{E}_t\hat{l}_{nt+1} \quad &\approx \quad \sum_{n}^{N}\left(\frac{\partial J}{\partial \log(\bar{\mu}_{int})}\mathbb{E}_t\hat{\mu}_{int} + \frac{\partial J}{\partial \log(\bar{l}_{nt})}\mathbb{E}_t\hat{l}_{nt}(z^{t-1})\right) \\
&= \quad \sum_{n}^{N}\frac{\bar{\mu}_{int}\bar{l}_{it}}{\bar{l}_{nt+1}}\Big(\mathbb{E}_t\hat{\mu}_{int} + \mathbb{E}_t\hat{l}_{nt}(z^{t-1})\Big)
\end{aligned} \tag{A.121}$$

**First-order Approximation to trade share and the price index**

Let $\log \lambda_{nit} = \log\left[z_{it}\left(\frac{w_{it}\kappa_{nit}}{P_{nt}}\right)^{-\theta}\right] \equiv K(w_{it}, \kappa_{nit}, P_{nt}, z_{nt})$. Then, a first-order approximation of $\mathbb{E}_t\hat{\lambda}_{nit}$ is:

$$\begin{aligned}
\mathbb{E}_t\hat{\lambda}_{nit} \quad &= \quad \left(\frac{\partial K}{\partial \log(\bar{w}_{it})}\mathbb{E}_t\hat{w}_{it} + \frac{\partial K}{\partial \log(\bar{\kappa}_{int})}\mathbb{E}_t\hat{\kappa}_{int} + \frac{\partial K}{\partial \log(\bar{P}_{nt})}\mathbb{E}_t\hat{P}_{nt} + \frac{\partial K}{\partial \log(\bar{z}_{it})}\hat{z}_{it}\right) \\
&= \quad -\theta(\mathbb{E}_t\hat{w}_{it} + \mathbb{E}_t\hat{\kappa}_{int} - \mathbb{E}_t\hat{P}_{nt}) + \hat{z}_{it}
\end{aligned} \tag{A.122}$$

Let $\log P_{nt} = \log\left(\sum_i z_{it}(w_{it}\kappa_{nit})^{-\theta}\right)^{-\frac{1}{\theta}} \equiv L(\mathbf{w_t}, \boldsymbol{\kappa_{nt}}, \mathbf{z_t})$ where $\boldsymbol{\kappa_{nt}} = [\kappa_{n1t}, ..., \kappa_{nNt}]'$. Then, a first-order approximation of $\mathbb{E}_t \hat{P}_{nt}$ is:

$$\begin{aligned}
\mathbb{E}_t \hat{P}_{nt} &\approx \sum_i \Big(\frac{\partial L}{\partial \log(\bar{w}_{it})}\mathbb{E}_t \hat{w}_{it} + \frac{\partial L}{\partial \log(\bar{\kappa}_{int})}\mathbb{E}_t \hat{\kappa}_{int} + \frac{\partial L}{\partial \log(\bar{z}_{it})}\hat{z}_{it}\Big) \\
&= \sum_i \bar{\lambda}_{nit}\Big(\mathbb{E}_t \hat{w}_{it} + \mathbb{E}_t \hat{\kappa}_{int} - \frac{1}{\theta}\hat{z}_{it}\Big) \qquad \text{(A.123)}
\end{aligned}$$

**First-order Approximation to labor market clearing**

Let $\log\Big(w_{nt}l_{nt}\Big) = \log\Big(\sum_i \lambda_{int}w_{it}l_{it}\Big) \equiv M(\boldsymbol{\lambda_{nt}}, \mathbf{w_t}, \mathbf{l_t})$ where $\boldsymbol{\lambda_{nt}} = [\lambda_{1nt}, ..., \lambda_{Nnt}]'$. Then, a first-order approximation of $\mathbb{E}_t \hat{w}_{it}$ is:

$$\begin{aligned}
\mathbb{E}_t \hat{w}_{nt} + \mathbb{E}_t \hat{l}_{nt} &\approx \sum_i \Big(\frac{\partial M}{\partial \log(\bar{\lambda}_{nit})}\hat{\lambda}_{int} + \frac{\partial M}{\partial \log(\bar{w}_{it})}\mathbb{E}_t \hat{w}_{it} + \frac{\partial M}{\partial \log(\bar{l}_{it})}\mathbb{E}_t \hat{l}_{it}\Big) \\
&= \sum_i \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{\bar{w}_{nt}\bar{l}_{nt}}\Big(\mathbb{E}_t \hat{\lambda}_{int} + \mathbb{E}_t \hat{w}_{it} + \mathbb{E}_t \hat{l}_{it}\Big) \qquad \text{(A.124)}
\end{aligned}$$

### A.6.2 Second-Order Derivatives

**Second-order approximation of value function**

For $\bar{x}, \bar{y} \in \{\log(\bar{w}_{nt+1}), \log(\bar{P}_{nt+1}), \bar{v}_{nt+2}\}_{n=1}^N$ and $\hat{x}, \hat{y} \in \left\{\hat{w}_{nt+1}, \hat{P}_{nt+1}, \hat{v}_{nt+2}\right\}_{n=1}^N$, a second-order approximation of expected values in equation (8) is given by

$$\begin{aligned}
\mathbb{E}_t \hat{v}_{nt+1} &\approx \frac{\partial F_n}{\partial \log(\bar{w}_{nt+1})}\mathbb{E}_t \hat{w}_{nt+1} + \frac{\partial F_n}{\partial \log(\bar{P}_{nt+1})}\mathbb{E}_t \hat{P}_{nt+1} + \frac{\partial F_n}{\partial \bar{\mathbf{v}}_{\mathbf{t+2}}}\hat{\mathbf{v}}_{\mathbf{t+2}} + \frac{1}{2}\sum_{x,y}\frac{\partial^2 F_n}{\partial \bar{x}\partial \bar{y}}\mathbb{E}_t \hat{x}\hat{y} \\
&= \frac{\partial F_n}{\partial \log(\bar{w}_{nt+1})}\mathbb{E}_t \hat{w}_{nt+1} + \frac{\partial F_n}{\partial \log(\bar{P}_{nt+1})}\mathbb{E}_t \hat{P}_{nt+1} + \frac{\partial^2 F_n}{\partial \log(\bar{w}_{nt+1})\partial \log(\bar{P}_{nt+1})}\mathbb{E}_t \hat{w}_{nt+1}\hat{P}_{nt+1} \\
&\quad + \frac{1}{2}\frac{\partial^2 F_n}{\partial \log(\bar{w}_{nt+1})^2}\mathbb{E}_t \hat{w}^2_{nt+1} + \frac{1}{2}\frac{\partial^2 F_n}{\partial \log(\bar{P}_{nt+1})^2}\mathbb{E}_t \hat{P}^2_{nt+1} \\
&\quad + \sum_k \frac{\partial F_n}{\partial \bar{v}_{kt+2}} \cdot \mathbb{E}_t \hat{v}_{kt+2} + \frac{1}{2}\sum_m \sum_k \frac{\partial^2 F_n}{\partial \bar{v}_{mt+2}\partial \bar{v}_{kt+2}} \cdot \mathbb{E}_t \hat{v}_{mt+2}\hat{v}_{kt+2} \qquad \text{(A.125)}
\end{aligned}$$

where

$$\frac{\partial F_n}{\partial \log(\bar{w}_{nt+1})} = \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{w}_{nt+1})}, \quad \frac{\partial F_n}{\partial \log(\bar{P}_{nt+1})} = \frac{\partial U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{P}_{nt+1})},$$
$$\frac{\partial^2 F_n}{\partial \log(\bar{w}_{nt+1})^2} = \frac{\partial^2 U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{w}_{nt+1})^2}, \quad \frac{\partial^2 F_n}{\partial \log(\bar{P}_{nt+1})^2} = \frac{\partial^2 U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{P}_{nt+1})^2}$$
$$\frac{\partial^2 F_n}{\partial \log(\bar{w}_{nt+1})\partial \log(\bar{P}_{nt+1})} = \frac{\partial^2 U(\bar{w}_{nt+1}, \bar{P}_{nt+1})}{\partial \log(\bar{w}_{nt+1})\partial \log(\bar{P}_{nt+1})}, \frac{\partial F_n}{\partial \bar{v}_{kt+2}} = \beta \bar{\mu}_{nkt+1},$$
$$\frac{\partial^2 F_i}{\partial \bar{v}_{kt+2}\partial \bar{v}_{mt+2}} = \beta \frac{\partial \bar{\mu}_{nmt+1}}{\partial \bar{v}_{kt+2}} = \frac{\beta^2}{\nu} \bar{\mu}_{nmt+1}\Big(\mathbb{I}(k=m) - \bar{\mu}_{nkt+1}\Big)$$

**Second-order approximation of mobility flows**

Let $\log \mu_{nit}(z^t) = \log \frac{\exp\left(\beta v_{it+1}(z^t)\right] - m_{nit}\Big)^{1/\nu}}{\sum_{h=1}^N \exp(\beta v_{ht+1}(z^t) - m_{nht})^{1/\nu}} \equiv G_{ni}(\mathbf{v_{t+1}}(z^t))$. A second-order approximation of expected values in equation (3) is given by

$$\mathbb{E}_t \hat{\mu}_{nit}(z^t) \approx \sum_k \frac{\partial G_{ni}}{\partial \bar{v}_{kt+1}} \mathbb{E}_t \hat{v}_{kt+1} + \frac{1}{2}\sum_{k,m} \frac{\partial^2 G_{ni}}{\partial \bar{v}_{kt+1}\partial \bar{v}_{mt+1}} \mathbb{E}_t \hat{v}_{kt+1}\hat{v}_{mt+1} \tag{A.126}$$

where

$$\begin{aligned}
\frac{\partial G_{ni}}{\partial \bar{v}_{kt+1}} &= \frac{\beta}{\nu}\Big(\mathbb{I}(k=i) - \bar{\mu}_{nkt}\Big) \\
\frac{\partial^2 G_{ni}}{\partial \bar{v}_{kt+1}\partial \bar{v}_{mt+1}} &= \frac{\beta}{\nu}\Big(\mathbb{I}(k=i)\frac{\partial \bar{\mu}_{nit}}{\partial \bar{v}_{mt+1}} - \bar{\mu}_{nkt}\frac{\partial \bar{\mu}_{nit}}{\partial \bar{v}_{mt+1}} - \cdot \bar{\mu}_{nit} \cdot \frac{\partial \bar{\mu}_{nkt}}{\partial \bar{v}_{mt+1}}\Big) \\
&= (\frac{\beta}{\nu})^2 \cdot \Big[\mathbb{I}(k=i)\bar{\mu}_{nit}(\mathbb{I}(m=i) - \bar{\mu}_{nmt}) - \bar{\mu}_{nkt}\bar{\mu}_{nit}[\mathbb{I}(m=i) - \bar{\mu}_{nmt} + \mathbb{I}(m=k) - \bar{\mu}_{nmt}]\Big].
\end{aligned}$$

**Second-order approximation of law of motion of labor**

Let $\log l_{nt+1}(z^t) = \log \sum_i^N \mu_{int}(z^t) l_{it}(z^{t-1}) \equiv J_n(\mathbf{v_{t+1}}(z^t), \mathbf{l_t}(z^{t-1}))$. For notational ease, let $\bar{\psi}_{int+1} = \frac{\bar{\mu}_{int}\bar{l}_{it}}{\bar{l}_{nt+1}}$. Then, a second-order approximation of $\mathbb{E}_t \hat{l}_{nt+1}$ is:

$$\begin{aligned}
\mathbb{E}_t \hat{l}_{nt+1}(z^t) \approx & \sum_{k=1}^N \frac{\partial J_n}{\partial \bar{v}_{kt+1}} \hat{v}_{kt+1} + \sum_{k=1}^N \frac{\partial J_n}{\partial ln(\bar{l}_{kt})} \cdot \mathbb{E}_t \hat{l}_{kt} \\
& + \frac{1}{2}\sum_{k,m} \frac{\partial^2 J_n}{\partial \bar{v}_{kt+1}\partial \bar{v}_{mt+1}} \mathbb{E}_t \hat{v}_{kt+1}\hat{v}_{mt+1} + \frac{1}{2}\sum_{k,m} \frac{\partial^2 J_n}{\partial ln(\bar{l}_{kt})\partial ln(\bar{l}_{mt})} \mathbb{E}_t \hat{l}_{kt}\hat{l}_{mt} \\
& + \sum_{k,m} \frac{\partial^2 J_n}{\partial \bar{v}_{kt+1}\partial ln(\bar{l}_{mt})} \mathbb{E}_t \hat{v}_{kt+1}\hat{l}_{mt},
\end{aligned} \tag{A.127}$$

where

$$
\begin{aligned}
\frac{\partial J_n}{\partial \bar{v}_{kt+1}} &= \frac{\beta}{\nu}\sum_{i=1}^{N}\bar{\psi}_{int+1}\Big(\mathbb{I}(k=n)-\bar{\mu}_{ikt}\Big)\\
\frac{\partial J_n}{\partial \log(\bar{l}_{kt})} &= \frac{\bar{\mu}_{knt}\bar{l}_{kt}}{\bar{l}_{nt+1}} \equiv \bar{\psi}_{knt+1}\\
\frac{\partial^2 J_n}{\partial \log(\bar{l}_{kt})\partial \log(\bar{l}_{mt})} &= \mathbf{1}(k=m)\bar{\psi}_{mnt+1}-\bar{\psi}_{knt+1}\bar{\psi}_{mnt+1}\\
\frac{\partial^2 J_n}{\partial \bar{v}_{kt+1}\partial \bar{v}_{mt+1}} &= \frac{\partial^2 J_n}{\partial \bar{v}_{mt+1}\partial \bar{v}_{kt+1}} = \frac{\partial \frac{\partial J_n}{\partial \bar{v}_{kt+1}}}{\partial \bar{v}_{mt+1}} = \frac{\partial \frac{\frac{\beta}{\nu}\sum_{i=1}^{N}\bar{l}_{it}\cdot\bar{\mu}_{int}\Big(\mathbb{I}(k=n)-\bar{\mu}_{ikt}\Big)}{\bar{l}_{nt+1}}}{\partial \bar{v}_{mt+1}}\\
&= (\frac{\beta}{\nu})^2\cdot\Big[-\sum_{i=1}^{N}\bar{\mu}_{ikt}\bar{\psi}_{int+1}[\mathbb{I}(m=n)+\mathbb{I}(m=k)-2\bar{\mu}_{imt}]\\
&\quad+\mathbb{I}(k=n)\sum_{i}\bar{\psi}_{int+1}(\mathbb{I}(m=n)-\bar{\mu}_{imt})\\
&\quad-\big[\sum_{i=1}^{N}\bar{\psi}_{int+1}\big(\mathbb{I}(k=n)-\bar{\mu}_{ikt}\big)\big]\cdot\big[\sum_{i}\bar{\psi}_{int+1}\big(\mathbb{I}(m=n)-\bar{\mu}_{imt}\big)\big]\Big]\\
\frac{\partial^2 J_n}{\partial \bar{v}_{kt+1}\partial \log(\bar{l}_{mt})} &= \frac{\beta}{\nu}\bar{\psi}_{mnt+1}\big(\mathbb{I}(k=n)-\bar{\mu}_{mkt}\big)-\frac{\beta}{\nu}\bar{\psi}_{mnt+1}\big[\sum_{i}\bar{\psi}_{int+1}\big(\mathbb{I}(k=n)-\bar{\mu}_{ikt}\big)\big].
\end{aligned}
$$

**Second-order approximation of trade share and the price index**
First order approximation for trade share $\boldsymbol{\lambda}$ is exact. Therefore,

$$
\begin{aligned}
\mathbb{E}_t\hat{\lambda}_{nit} &= \Big(\frac{\partial K}{\partial \log(\bar{w}_{it})}\mathbb{E}_t\hat{w}_{it+1}+\frac{\partial K}{\partial \log(\bar{\kappa}_{int})}\mathbb{E}_t\hat{\kappa}_{int}+\frac{\partial K}{\partial \log(\bar{P}_{nt})}\mathbb{E}_t\hat{P}_{nt}+\frac{\partial K}{\partial \log(\bar{z}_{nt})}\mathbb{E}_t\hat{z}_{nt}\Big)\\
&= -\theta(\mathbb{E}_t\hat{w}_{it}+\mathbb{E}_t\hat{\kappa}_{int}-\mathbb{E}_t\hat{P}_{nt})+\mathbb{E}_t\hat{z}_{it} \qquad \text{(A.128)}
\end{aligned}
$$

Let $\log P_{nt} = \log\Big(\sum_i (w_{it}\kappa_{nit})^{-\theta}z_{it}\Big)^{-\frac{1}{\theta}} \equiv L(\mathbf{w_t},\boldsymbol{\kappa_{nt}},\mathbf{z_t})$. Then, a second-order approximation of $\mathbb{E}_t\hat{P}_{nt}$ is:

$$
\begin{aligned}
\mathbb{E}_t\hat{P}_{nt} \approx\; & \sum_{i=1}^{N}\frac{\partial L}{\partial ln(\bar{w}_{it})}\mathbb{E}_t\hat{w}_{it}+\sum_{i=1}^{N}\frac{\partial L}{\partial ln(\bar{\kappa}_{nit})}\mathbb{E}_t\hat{\kappa}_{nit}+\sum_{i=1}^{N}\frac{\partial L}{\partial ln(\bar{z}_{it})}\mathbb{E}_t\hat{z}_{it}\\
+\; & \sum_{i,m}\frac{1}{2}\frac{\partial^2 L}{\partial ln(\bar{w}_{it})\partial ln(\bar{w}_{mt})}\mathbb{E}_t\hat{w}_{it}\hat{w}_{mt}+\sum_{i,m}\frac{1}{2}\frac{\partial^2 L}{\partial ln(\bar{\kappa}_{nit})\partial ln(\bar{\kappa}_{nmt})}\mathbb{E}_t\hat{\kappa}_{nit}\hat{\kappa}_{nmt}+\sum_{i,m}\frac{1}{2}\frac{\partial^2 L}{\partial ln(\bar{z}_{it})\partial ln(\bar{z}_{mt})}\mathbb{E}_t\hat{z}_{it}\hat{z}_{mt}\\
+\; & \sum_{i,m}\frac{\partial^2 L}{\partial ln(\bar{w}_{it})\partial ln(\bar{\kappa}_{nmt})}\mathbb{E}_t\hat{w}_{it}\hat{\kappa}_{nmt}+\sum_{i,m}\frac{\partial^2 L}{\partial ln(\bar{w}_{it})\partial ln(\bar{z}_{mt})}\mathbb{E}_t\hat{w}_{it}\hat{z}_{mt}+\sum_{i,m}\frac{\partial^2 L}{\partial ln(\bar{\kappa}_{nit})\partial ln(\bar{z}_{mt})}\mathbb{E}_t\hat{\kappa}_{nit}\hat{z}_{mt}
\end{aligned}
$$

(A.129)

where

$$\frac{\partial L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{w}_{it})} = \bar{\lambda}_{nit}, \frac{\partial L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{\kappa}_{nit})} = \bar{\lambda}_{nit}, \quad \frac{\partial L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{z}_{nt})} = -\frac{1}{\theta}\bar{\lambda}_{nit}$$

$$\frac{\partial^2 L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{w}_{it})\partial \log(\bar{w}_{mt})} = \frac{\partial \bar{\lambda}_{nit}}{\partial \log(\bar{w}_{mt})} = \mathbf{I}(i=m)(-\theta)\bar{\lambda}_{nit} + \theta\bar{\lambda}_{nit}\bar{\lambda}_{nmt}$$

$$\frac{\partial^2 L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{\kappa}_{nit})\partial \log(\bar{\kappa}_{nmt})} = \frac{\partial \bar{\lambda}_{nit}}{\partial \log(\bar{\kappa}_{nmt})} = \mathbf{I}(i=m)(-\theta)\bar{\lambda}_{nit} + \theta\bar{\lambda}_{nit}\bar{\lambda}_{nmt}$$

$$\frac{\partial^2 L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{z}_{it})\partial \log(\bar{z}_{mt})} = -\frac{1}{\theta}\mathbf{I}(i=m)\bar{\lambda}_{nmt} + \left(\frac{1}{\theta}\right)\bar{\lambda}_{nit}\bar{\lambda}_{nmt}$$

$$\frac{\partial^2 L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{w}_{it})\partial \log(\bar{\kappa}_{nmt})} = \frac{\partial \bar{\lambda}_{nit}}{\partial \log(\bar{\kappa}_{nmt})} = \mathbf{I}(i=m)(-\theta)\bar{\lambda}_{nit} + \theta\bar{\lambda}_{nit}\bar{\lambda}_{nmt}$$

$$\frac{\partial^2 L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{w}_{it})\partial \log(\bar{z}_{mt})} = \frac{\partial \bar{\lambda}_{nit}}{\partial \log(\bar{z}_{mt})} = \mathbf{I}(m=i)\bar{\lambda}_{nmt} - \bar{\lambda}_{nit}\bar{\lambda}_{nmt}$$

$$\frac{\partial^2 L(\mathbf{w},\boldsymbol{\kappa},\mathbf{z})}{\partial \log(\bar{\kappa}_{nit})\partial \log(\bar{z}_{mt})} = \frac{\partial \bar{\lambda}_{nit}}{\partial \log(\bar{z}_{mt})} = \mathbf{I}(m=i)\bar{\lambda}_{nmt} - \bar{\lambda}_{nit}\bar{\lambda}_{nmt}$$

**Second-order approximation of labor market clearing**

Let $\log\Big(w_{nt}l_{nt}\Big) = \log\Big(\sum_i \lambda_{int}w_{it}l_{it}\Big) \equiv M(\boldsymbol{\lambda_{nt}},\mathbf{w_t},\mathbf{l_t})$. Then, a second-order approximation of $\mathbb{E}_t\hat{w}_{nt} + \mathbb{E}_t\hat{l}_{nt}$ is:

$$\begin{aligned}
\mathbb{E}_t\hat{w}_{nt} + \mathbb{E}_t\hat{l}_{nt} \approx \quad & \sum_{i=1}^{N}\frac{\partial M}{\partial ln(\bar{\lambda}_{int})}\cdot\mathbb{E}_t\hat{\lambda}_{int} + \sum_{i=1}^{N}\frac{\partial M}{\partial ln(\bar{w}_{it})}\cdot\mathbb{E}_t\hat{w}_{it} + \sum_{i=1}^{N}\frac{\partial M}{\partial ln(\bar{l}_{it})}\cdot\mathbb{E}_t\hat{l}_{it} \\
& + \sum_{i,m}\frac{1}{2}\frac{\partial^2 M}{\partial ln(\bar{w}_{it})\partial ln(\bar{w}_{mt})}\cdot\mathbb{E}_t\hat{w}_{it}\hat{w}_{mt} + \sum_{i,m}\frac{1}{2}\frac{\partial^2 M}{\partial ln(\bar{\lambda}_{int})\partial ln(\bar{\lambda}_{mnt})}\cdot\mathbb{E}_t\hat{\lambda}_{int}\hat{\lambda}_{mnt} \\
& + \sum_{i,m}\frac{1}{2}\frac{\partial^2 M}{\partial ln(\bar{l}_{it})\partial ln(\bar{l}_{mt})}\cdot\mathbb{E}_t\hat{l}_{it}\hat{l}_{mt} \\
& + \sum_{i,m}\frac{\partial^2 M}{\partial ln(\bar{\lambda}_{int})\partial ln(\bar{w}_{mt})}\cdot\mathbb{E}_t\hat{\lambda}_{int}\hat{w}_{mt} + \sum_{i,m}\frac{\partial^2 M}{\partial ln(\bar{\lambda}_{int})\partial ln(\bar{l}_{mt})}\cdot\mathbb{E}_t\hat{\lambda}_{int}\hat{l}_{mt} \\
& + \sum_{i,m}\frac{\partial^2 M}{\partial ln(\bar{w}_{it})\partial ln(\bar{l}_{mt})}\cdot\mathbb{E}_t\hat{w}_{it}\hat{l}_{mt}
\end{aligned} \tag{A.130}$$

where

$$\frac{\partial M}{\partial ln(\bar{\lambda}_{int})} = \frac{\partial M}{\partial ln(\bar{w}_{it})} = \frac{\partial M}{\partial ln(\bar{l}_{it})} = \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{\bar{w}_{nt}\bar{l}_{nt}}$$

$$\frac{\partial^2 M}{\partial ln(\bar{w}_{it})\partial ln(\bar{w}_{mt})} = 1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})} - \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}$$

$$\frac{\partial^2 M}{\partial ln(\bar{\lambda}_{int})\partial ln(\bar{\lambda}_{mnt})} = 1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})} - \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}$$

$$\frac{\partial^2 M}{\partial ln(\bar{l}_{it})\partial ln(\bar{l}_{mt})} = 1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})} - \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}$$

$$\frac{\partial^2 M}{\partial ln(\bar{\lambda}_{int})\partial ln(\bar{w}_{mt})} = 1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})} - \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}$$

$$\frac{\partial^2 M}{\partial ln(\bar{\lambda}_{int})\partial ln(\bar{l}_{mt})} = 1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})} - \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}$$

$$\frac{\partial^2 M}{\partial ln(\bar{w}_{it})\partial ln(\bar{l}_{mt})} = 1(m=i)\frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}}{(\bar{w}_{nt}\bar{l}_{nt})} - \frac{\bar{\lambda}_{int}\bar{w}_{it}\bar{l}_{it}\bar{\lambda}_{mnt}\bar{w}_{mt}\bar{l}_{mt}}{(\bar{w}_{nt}\bar{l}_{nt})^2}$$

# Appendix B Application

**List of countries and regions.** We calibrate the model to the 50 U.S. states and 28 other regions including the rest of the world. The countries in addition to the United States are Australia, Austria, Balkans (aggregating Bulgaria, Croatia, and Greece), Baltic States (aggregating Estonia, Latvia, and Lithuania), Benelux (Belgium, Netherlands, Luxembourg), Brazil, Canada, Switzerland, China, Central Europe (aggregating Poland, Czech Republic, Slovakia, Slovenia, and Hungary), Germany, Spain, France, United Kingdom, Indonesia, India, Ireland, Italy, Japan, South Korea, Mexico, Nordic Countries (aggregating Denmark, Finland, Norway, and Sweden), Portugal, Romania, Russia, Turkey, Taiwan, and the rest of world. Regarding the 50 U.S. States, District of Columbia is aggregated into Virginia.

**Definition of the five broad U.S. regions.** In the climate change application (Section 3), our quantification is at the state level, but some results are reported by five broad U.S. regions: Northeast (NE), Southeast (SE), Midwest (ME), Southwest (SW), and West (WE). These regions are defined as:

- Northeast: Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont.
- Southeast: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Tennessee, South Carolina, Virginia, and West Virginia.
- Midwest: Iowa, Illinois, Indiana, Kansas, Michigan, Missouri, Ohio, Wisconsin, Minnesota, Nebraska, North Dakota, and South Dakota.
- Southwest: Arizona, New Mexico, Oklahoma, and Texas.

- West: Alaska, California, Colorado, Hawaii, Idaho, Montana, Nevada, Oregon, Utah, Washington, and Wyoming.

**List of sectors.** In the climate change application we calibrate the model to 3 sectors; agriculture, manufacturing, and service. Table B.1 summarizes how each sector in the datasets we use is aggregated into three sectors. In the China shock application, we calibrate the model to 2 sectors; manufacturing and others. We aggregate agriculture and service sectors defined in table B.1 to a non-manufacturing sector.

Table B.1: List of sectors and dataset

| Dataset | Agriculture | Manufacturing | Service |
|---|---|---|---|
| WIOD | Agriculture, Forestry, Fishing, Mining<br>(sectors A and B) | Manufacturing<br>(sector C) | The rest |
| BEA:GDP | Agriculture, forestry, fishing, and hunting<br>Mining, quarrying, and oil and gas extraction | Manufacturing | The rest |
| BEA:EMPLOYMENT | Farm employment (linecode 70)<br>Forestry, fishing, and related activities<br>Mining, quarrying, and oil and gas extraction | Manufacturing<br>(linecode 500) | The rest |
| CENSUS | 111 Agricultural Products<br>112 Livestock & Livestock Products<br>113 Forestry Products, Nesoi<br>114 Fish, Fresh/chilled/frozen<br>& Other Marine Products<br>211 Oil & Gas<br>212 Minerals & Ores | The rest | N/A |
| CFS | N/A | Manufacturing<br>(sectors 1-12) | N/A |
| LEHD | Agriculture, forestry, fishing, and hunting<br>Mining, quarrying, and oil and gas extraction | Manufacturing<br>(sectors 31-33) | The rest |

**Climate Belief Data.** We use the 2014 survey of Yale Climate Opinion data at the US-state level. The share of believers in each state is defined by the share of "Yes" responses to the question: "What do you think: Do you think that global warming is happening?" We consider the respondents who answered "No" or "Don't know" as climate skeptics. The average share of believers across the states in 2014 is 62.6% and the standard deviation across states is 5.32.

**Temperature Projection Data.** The temperature projection data we use is from Climate Impact Lab. Climate Impact Lab provides six cases: two scenarios (RCP 4.5 & RCP 8.5) with three paths for temperature rise for each scenarios (median, 5th percentile, 95the percentile) of future temperature. We adopt RCP 4.5 median case and use the 1-in-20 low and 1-in-20 high values to compute the variance $(\sigma_n^{t,t'})^2$ of the temperature projection.

**Constructing migration matrix for China Shock.** As mentioned in Section 4.1, LEHD job-to-job flow data miss some states over 2000-2009. Table B.2 shows the list of missing states in LEHD. We fill in the missing values using migration in Caliendo et al. (2019), who incorporate 23 sectors including the non-employment sector. We neglect migration in & out of the non-employment

Table B.2: List of Missing States in LEHD

| Year | States |
|---|---|
| 2000 | AL, AZ, AR, KY, MA, MS, NH, WY |
| 2001-2002 | AZ, AR, MA, MS, NH, MI |
| 2003 | AR, MA |
| 2004-2009 | MA |

sector and normalize the rest of the values. Then we aggregate the remaining sectors into 2 sectors; sectors 2-13 and sectors 14-23 in Caliendo et al. (2019) are mapped to the manufacturing and non-manufacturing sectors in our model, respectively.

**Model inversion using Penn World Table data.** Here we explain how we measure the variables when we invert our model to recover the productivity parameter of country $n$ in sector $j$, $z_{nj}^t$:

$$\pi_{nn,j}^t = z_{nj}^t \left( \frac{w_{nj}^t}{P_{nj}^t} \right)^{-\theta \cdot \gamma_{nj}}.$$

Trade share $\pi_{nn,j}^t$ is obtained from EORA. $\gamma_{nj}$, the input shares of intermediate inputs, come from WIOD 2007 data. For wage and price, we use Penn World Table version 9.1. We write the corresponding variable names in PWT dataset inside each parenthesis. Wage in each region is measured by the normalized GDP multiplied by the share of labor compensation in GDP (variable *labsh* in Penn World Table) and divided by the total number of employment (variable *emp*):

$$wage_n^t = GDP_n^t \cdot \frac{Laborshare_n^t}{Employment_n^t},$$

where $GDP_n^t$ is output-side real GDP at period $t$ multiplied by the currency rate (variables *cgdpo* and *pl_gdpo*). For manufacturing price index, $P_{n1}^t$, we average the import and export prices (variables *pl_m* and *pl_x*). The price index for the non-manufacturing sector, $P_{n2}^t$, is measured by the price level of household consumption (variable *pl_c*).