Optimal Dynamic Spatial Policy [*]

Eric Donald[†] Masao Fukui[‡] Yuhei Miyauchi[§]
University of Pittsburgh *Boston University* *Boston University*

August 2025

Abstract

We study the optimal allocation of population and consumption in a dynamic spatial general equilibrium model with frictional migration, where households' idiosyncratic location preference shocks are private information. We derive a recursive formula for the constrained-efficient allocation, capturing the trade-off between consumption smoothing and efficient migration. In a quantitative model calibrated to the US economy featuring both cross-state migration and risk-free savings, we find that the constrained-efficient allocation features *lower* population but *higher* average consumption in less productive states than the status quo, achieving efficiency and spatial redistribution simultaneously through dynamic incentives. In response to local negative productivity shocks, the constrained-efficient allocation features more front-loaded consumption profiles than the status quo, with systematic heterogeneity linked to the location's pre-shock fundamentals.

[*]We thank Costas Arkolakis, Lorenzo Caliendo, Jonathan Dingel, Greg Howard, David Lagakos, Michael Peters, and Pascual Restrepo, as well as seminar participants at Barcelona Summer Forum, Boston University, Green Line Macro Meeting, SED Annual Meeting, University of Pittsburgh, University of Tokyo, and Yale for their helpful comments.

[†]University of Pittsburgh (e-mail: eric.donald@pitt.edu).
[‡]Boston University (e-mail: mfukui@bu.edu).
[§]Boston University (e-mail: miyauchi@bu.edu).

1 Introduction

What is the optimal spatial distribution of population? How does it differ from the observed equilibrium? And what policies can improve aggregate welfare? These questions lie at the center of ongoing academic and policy debates surrounding the design of place-based interventions (Fajgelbaum and Gaubert 2025). However, much of the existing literature addresses them in static frameworks, overlooking the reality that migration is gradual and forward-looking (Caliendo, Dvorkin, and Parro 2019). The focus on static settings also limits the ability to speak about the optimal response to regional shocks. In this paper, we examine these questions through a dynamic perspective. We show, both theoretically and quantitatively, that accounting for migration dynamics fundamentally alters the design and consequences of optimal place-based policies.

We consider a general environment with many heterogeneous locations that differ in productivity, amenities, trade costs, and agglomeration externalities, which potentially evolve over time. In each period, households draw idiosyncratic location preference shocks and choose their residential location for the following period. The primary objective of our paper is to characterize the socially optimal allocation of consumption and migration in this environment.

The key constraint for the planner is that idiosyncratic location preference shocks are private information. If the shocks were observable, the planner could fully smooth consumption by equalizing the marginal utility of income across locations and time, and assign each household's location contingent on their preference shocks to maximize *social* surplus. However, when shocks are unobservable, such allocations are not incentive compatible, as households would benefit by choosing locations that maximize *private* surplus. Therefore, the planner must design allocations that respect incentive compatibility for migration decisions. We refer to the resulting allocation as the "constrained-efficient allocation", as this allocation maximizes welfare subject to incentive compatibility constraints, independent of the market structure or specific policy instruments available to the planner.

The main challenge in solving the constrained-efficient allocation arises from the high dimensionality of the state space and the choice variables. The planner must track the full distribution of agents by migration history and assign consumption and migration accordingly. We show that this complex problem can be decomposed into tractable component problems: for each agent, the planner determines consumption and location assignments conditional on their current location and promised utility, which summarizes the household's entire migration history. Because each component problem depends only on each agent's current location and promised utility, this structure significantly reduces the dimensionality of the problem, yielding both analytical and computational tractability.

Our first main theoretical result is a recursive formula that characterizes the constrained-efficient allocation. This formula summarizes the planner's central trade-off: incentivizing migration toward more productive (i.e. higher net social surplus) regions while preserving consumption smoothing. To encourage relocation to productive areas, the planner can offer either higher current consumption or promise higher future consumption conditional on staying in those areas. At the same time, the planner values smoothing consumption for individuals who, due to idiosyncratic shocks, remain in less productive regions where marginal utility is higher. The resulting optimal allocation is inherently dynamic: individuals who land in unproductive regions receive high initial consumption that gradually declines if they remain, while those who move to more productive regions are rewarded with increases in consumption.

Our results indicate that the provision of dynamic incentives plays a crucial role in achieving the constrained-efficient allocation. These incentives are delivered by tailoring consumption to households' migration histories, making the optimal policy inherently "history-dependent". In practice, such policies may involve providing temporary subsidies to residents in declining regions or offering incentives for long-term residence in more productive areas. Our analysis characterizes the best possible outcomes attainable under such history-dependent policy space.

At the same time, policymakers may face informational or administrative constraints that limit their ability to condition policies on households' migration histories. For this reason, we also consider a "history-independent constrained-efficient allocation", in which consumption depends only on an agent's current location. We derive an analogous recursive formula for this case that summarizes the trade-off between consumption smoothing and efficient migration. However, by removing the ability to condition consumption on past migration decisions, the planner has limited ability to use dynamic incentives. As a result, the trade-off between encouraging migration to productive regions and smoothing consumption becomes more pronounced.

Building on these theoretical results, we quantify how the constrained-efficient allocation differs from that of the observed equilibrium. To this end, we apply our model to US states, with households making forward-looking migration decisions (Caliendo et al. 2019) and choosing consumption and savings using risk-free assets subject to occasionally binding borrowing constraints (Bewley 1986, Huggett 1993, Aiyagari 1994, Imrohoroğlu 1989). Bilateral migration and trade costs across states are disciplined using 2017 data on migration and trade flows, under the assumption that the economy is in steady-state in that year. We further calibrate households' ability to smooth consumption by targeting empirical estimates of the marginal propensity to consume.

A key challenge in the computation of the constrained-efficient allocation and status quo equilibrium is the high dimensionality of the state space and control variables. The computation of the steady-state requires finding fixed points in all locations' prices or Lagrange multipliers, as well as the population size distribution, where we need to repeatedly solve Bellman equations in the inner loop. Since the Bellman equations involve high-dimensional optimization, naive algorithms such as grid search or Newton's method are practically infeasible. To address this challenge, we extend the endogenous grid point method of Carroll (2006), which speeds up value function iteration by orders of magnitude, making the steady-state computation feasible.

Equipped with our calibrated model and computational methodology, we first compare the steady-state of the status quo equilibrium with that of the constrained-efficient allocation, finding substantial and systematic differences between the two allocations. The constrained-efficient allocation features more population in states with higher output per capita (e.g. Washington) and less population in states with lower output per capita (e.g. Mississippi). Interestingly, despite this population reallocation, average consumption per capita is *lower* in the former states and *higher* in the latter states, resulting in more equal consumption across states than in the status quo. This pattern reflects the role of dynamic incentives. In unproductive states, consumption is front-loaded. This threat to cut consumption in the future incentivizes households to out-migrate without cutting consumption today. In this way, dynamic incentives serve to promote both spatial equity and efficiency. This is in stark contrast to static settings, where the two objectives typically conflict (Gaubert, Kline, and Yagan 2021, Ales and Sleet 2022, Donald, Fukui, and Miyauchi 2025).

We next examine the history-independent constrained-efficient allocation, in which consumption only depends on current location. This allocation also shifts population toward more productive states, but it yields *higher* consumption in such states and *lower* consumption in less productive ones, producing greater inequality across locations than in the status quo. These patterns reflect the planner's limited ability to use dynamic incentives when consumption can only depend on current location but not past migration decisions. Therefore, the planner faces a sharper trade-off between efficient migration and spatial redistribution, analogous to the efficiency-equity trade-off emphasized in static models.

In the final part of the paper, we study transition dynamics in response to aggregate regional productivity shocks. A key challenge in studying aggregate shocks in our environment is again the high dimensionality of the state space: with aggregate shocks, one must track the full distribution of population over location and assets (in the status quo) or location and promised utility (in the constrained-efficient allocation). To address this challenge, we focus on a one-time shock that occurs with vanishingly small probability.[1] This setup allows us to compute a first-order approximation of the transition path using only the sequence-space Jacobian (Auclert, Bardóczy,

[1]This assumption is distinct from the commonly used "MIT shock" (one-time *unanticipated* shock). "MIT shocks" are ill-suited to study optimal policy due to time inconsistency problems. In response to an unanticipated shock, the planner needs to re-optimize. However, once migration decisions are realized, the planner no longer has an incentive to maintain any ex-ante plan of consumption front-loading or back-loading. This re-optimization creates artificial dynamics unrelated to the aggregate shock itself. Similar approaches appear in the context of optimal risk-sharing contracts (Fukui 2020) and endogenous portfolio choice (Auclert, Rognlie, Straub, and Tapak 2024).

Rognlie, and Straub 2021) with respect to aggregate variables such as local wages and population size – significantly lower-dimensional objects.

We find that dynamic incentives play a central role in shaping the transition dynamics of the constrained-efficient allocation in response to localized negative productivity shocks. The planner aims to achieve two objectives simultaneously: incentivizing households to leave the adversely affected location and insuring them against the shock. These goals are accomplished by front-loading consumption: the planner first keeps the consumption drop minimal, and then gradually decreases it over time. As a result, households in negatively affected areas are insured but are strongly motivated to migrate before consumption declines. We show that the degree of front-loading differs systematically with the initial productivity of the shocked location, with more productive states featuring more front-loading. Such heterogeneity also leads to heterogeneity in the speed of transition. When a more productive location is hit by the shock, the transitions are slower relative to the status quo.

Overall, our results highlight the crucial role of dynamics in shaping optimal place-based policies, both in the steady-state and during transitions in response to aggregate shocks.

Related Literature. First, we contribute to the literature on dynamic spatial general equilibrium models. These frameworks have been used for various applications, such as regional incidence of import competition, the rise of automation, immigration shocks, or climate change (see Desmet and Parro (2025) for a recent survey). Early contributions focused on environments where households make dynamic migration decisions but are hand-to-mouth (e.g. Caliendo et al. 2019). More recent work incorporates forward-looking agents who make both migration and consumption-saving decisions (e.g. Giannone, Li, Paixão, and Pang 2023, Dvorkin 2023, Greaney 2023, Greaney, Parkhomenko, and Van Nieuwerburgh 2025).[2] Despite the increasing use of these models, little is known about the optimal allocation of population and consumption and how it differs from the laissez-faire outcome. Our contribution is to fill this gap both theoretically and quantitatively.

Several recent papers share the broad goal of analyzing optimal allocations and policies in dynamic spatial general equilibrium models, though they differ in focus and methodology. O'Connor (2024) examines optimal spatial transfers in response to regional productivity shocks to address dynamic inefficiencies arising from wage rigidity. He derives an analytical formula in a two-period model to correct for these externalities and quantitatively implements it a linearized, infinite-horizon environment where there is no role for policy in the steady-state. Lhuillier (2023)

[2]Other extensions consider hand-to-mouth migrants alongside forward-looking capitalists who make dynamic investment decisions (e.g. Kleinman, Liu, and Redding 2023, Bilal and Rossi-Hansberg 2023), or settings with dynamic agglomeration externalities (e.g. Allen and Donaldson 2020, Peters 2022). We show in Section 5 that our framework accommodates these environments as well.

explores a dynamic spatial model with learning externalities, comparing equilibrium and optimal allocations in a stylized setting. Our paper is different, both in focus and methods. Rather than focusing solely on (specific sources of) externalities, we study the trade-off between such efficiency considerations and consumption smoothing motives in a general class of fully-dynamic infinite-horizon model, both analytically and quantitatively. Analytically, we derive a recursive formula that characterizes this trade-off (inclusive of externalities) in the constrained-efficient allocation. Quantitatively, we solve for the constrained-efficient allocation in both the steady-state and the transition.

Our work is also related to Kurnaz, Michelini, Özdenören, and Sleet (2023), who study optimal taxation in a stationary dynamic discrete choice model, focusing on maximizing long-run welfare where location-specific consumption is constant over time. In contrast, we solve the full dynamic planning problem and derive a recursive formula for the constrained-efficient allocation, explicitly considering the dynamic path of aggregate fundamentals as well as each household's location choice history.

As noted earlier, our paper contributes more broadly to the literature on the optimal design of place-based policies in static settings (see Fajgelbaum and Gaubert (2025) for a review) by highlighting the role of dynamics. In this static literature, our paper extends the analysis of Gaubert et al. (2021), Ales and Sleet (2022), Guerreiro, Rebelo, and Teles (2023), Mongey and Waugh (2024), and Donald et al. (2025), who highlight the efficiency-equity trade-off. We extend these frameworks to a dynamic environment, incorporating forward-looking behavior and intertemporal trade-offs.

We also contribute to the macroeconomics and public finance literature that characterizes constrained-efficient allocations in dynamic general equilibrium models under information constraints, where planners must address moral hazard or adverse selection problems. Our recursive formulation builds on the work of Atkeson and Lucas (1992) and Farhi and Werning (2007), who examine optimal social insurance when idiosyncratic shocks are privately observed, as well as Hopenhayn and Nicolini (1997) and Veracierto (2022), who study optimal unemployment insurance under private search effort. We extend these frameworks to a dynamic discrete choice environment with many heterogeneous discrete options. To address the associated high-dimensional state space, we develop a computational approach and apply it to study the optimal allocation of population and consumption across US states.

2 Model Environment

We consider an economy consisting of J locations. Time is discrete and the horizon is infinite, $t = 0, 1, \ldots, \infty$. There is a unit measure of household dynasty indexed by $h \in [0, 1]$, each of which is endowed with a unit of labor in each period. In each period, consumption and production take place; households stochastically die and are replaced by newborns; and households decide where to live in the next period, given expected values, migration costs, and their idiosyncratic preference shocks. Locations differ in their productivities, amenities, trade costs, and agglomeration externalities, which potentially evolve over time. This section defines the model environment. We introduce specific market structures later when we discuss implementations and quantification. There is no aggregate uncertainty until Section 7.

Preferences. At the beginning of the period t, household h living in location j consumes a location-specific final good aggregator, denoted by $C_{jt}(h)$. This household's flow utility from their consumption in location j at time t is given by $u_{jt}(C_{jt}(h))$. We describe the associated production technology for final goods later.

After consumption and production take place, households die with probability $1-\omega$. Whenever households die, they are replaced by newborns in the same location. The newborn inherits the household index h, continuing the household dynasty. For brevity, we henceforth refer to these household dynasties simply as households.[3]

Finally, at the end of the period t, all households (including newborns) in location j draw a vector of idiosyncratic preference shocks and decide where to migrate. The location preference shocks for each destination k, $\epsilon_{jt} \equiv [\epsilon_{jkt}]_k$, are additively separable from $u_{jt}(C_{jt}(h))$. These preference shocks $[\epsilon_{jkt}]_k$ capture factors determining migration decisions that are specific to households and difficult to observe (e.g. finding a new job or relatives living in town). We do not impose particular assumptions on the distribution of preferences $G_{jt}(\epsilon_{jt})$, such as the independence across alternative options k or an extreme-value distribution, as commonly assumed in the literature (e.g. Artuç, Chaudhuri, and McLaren (2010), Kennan and Walker (2011), and Caliendo et al. (2019)). The only restriction is that shocks are independent over time.[4] Notice also that the mean of ϵ_{jkt} can arbitrarily depend on origin j, destination k, and time t, which captures migration utility costs that depend on location pairs and time. Throughout, we denote $\mathbb{E}_{jt}$ as the expectation operator over ϵ_{jt}, i.e. $\mathbb{E}_{jt}[x] \equiv \int x([\epsilon_{jkt}]_k) dG_{jt}([\epsilon_{jkt}]_k)$.

After observing their preference shocks, all households (including newborns) choose their migration destinations, with staying in their current location as one of the options. All households

[3] A special case of infinitely lived households (without overlapping generations) corresponds to $\omega = 1$. For our quantitative analysis, we focus on the case with $\omega < 1$. As in other environments, constrained-efficient allocations with private information feature immiseration in the long run whenever $\omega = 1$, implying that a steady-state distribution does not exist (Atkeson and Lucas 1992, Farhi and Werning 2007, Bloedel, Krishna, and Leukhina 2025).

[4] Since $G_{jt}(\epsilon_{jt})$ can flexibly depend on origin location j, our model accommodates a time dependence of idiosyncratic shocks through past location decisions. We rule out an explicit correlation of these shocks over time (e.g. Howard and Shao (2022)) to avoid introducing a state variable unobserved to the planner. See also Section 5.2 for an extension with multiple ex-ante heterogeneous household types observable to the planner.

discount the future with a discount factor $\beta \in (0, 1)$. Let $\ell_t(h) \in \{1, \dots, J\}$ be the location that household h lives at time t. The value function of household h living in location j at time t is recursively given by

$$v_{jt}(h) = u_{jt}(C_{jt}(h)) + \beta\omega\mathbb{E}_{jt}\left[\sum_k \mathbb{I}(\ell_{t+1}(h) = k)\{v_{kt+1}(h) + \epsilon_{jkt}\}\right]. \tag{1}$$

The value function of newborns h born in location j at the end of period t is

$$v^n_{jt}(h) = \beta\mathbb{E}_{jt}\left[\sum_k \mathbb{I}(\ell_{t+1}(h) = k)\{v_{kt+1}(h) + \epsilon_{jkt}\}\right], \tag{2}$$

where the superscript n denotes variables associated with newborns. This expression closely parallels the value function for surviving households in (1), with two differences: (i) newborns do not consume in period t, as they are born after consumption has occurred, and (ii) the effective discount factor is β rather than $\beta\omega$, as newborns do not die in the same period as they are born.

Technology. The location-specific final good aggregator is a composite of various goods, some of which can be tradable (e.g. food or manufacturing goods) or nontradable (e.g. housing or nontradable services). Instead of modeling each of these goods, we follow Adão, Costinot, and Donaldson (2017) and specify a reduced-form production technology for location-specific final good aggregators using labor services from various locations.[5] Specifically, non-tradable goods in each location j are produced according to the following production technology:

$$Y_{jt} = f_{jt}\left(\{l_{kjt}\}_k, \{L_{kt}\}_k\right), \tag{3}$$

where l_{kjt} is the labor service in location k used to produce the final products in j at time t, and L_{kt} is the total population size of location k. The labor inputs $\{l_{kjt}\}_k$ are direct production factors, while the total populations $\{L_{kt}\}_k$ capture agglomeration and congestion externalities that influence productivity in j as a function of local or nearby population levels.[6] We assume f_{jt} is constant returns to scale in $\{l_{kjt}\}_k$.

[5] As demonstrated by Adão et al. (2017), the above specification accommodates any single-factor neoclassical trade environments, such as Ricardian trade models as in Eaton and Kortum (2002) or an arbitrary form of input-output linkages as in Caliendo and Parro (2015). Specification (3) also accommodates endogenous amenities by reinterpreting some final consumption goods as amenities that can be either produced or influenced by the population size distribution.

[6] Section 5.3 shows that our analysis extends straightforwardly to an environment with dynamic agglomeration and congestion externalities from lagged population.

Resource Constraints. The goods market clearing conditions are

$$\int_0^1 C_{jt}(h)dh = f_{jt}\left(\{l_{kjt}\}_k, \{L_{kt}\}_k\right), \tag{4}$$

and the labor market clearing conditions require that population size in location j equals the demand for labor services such that

$$L_{jt} = \int_0^1 \mathbb{I}\left[\ell_t(h) = j\right] dh = \sum_k l_{jkt}. \tag{5}$$

3 First-Best Allocation

We begin by analyzing the first-best allocation, where the planner can perfectly observe the entire history of households' idiosyncratic location preference shocks, $\epsilon^t \equiv (\epsilon_0, \ldots, \epsilon_t)$, and tailor the consumption and migration allocation accordingly. Although this assumption is clearly unrealistic, the first-best serves as a useful stepping stone for understanding the constrained-efficient allocation – our main focus – which we study next in Section 4, for two reasons. First, it highlights the fundamental goals that the planner aims to achieve. Second, it allows us to introduce our methodology for reducing the dimensionality of the planning problem transparently, which will carry over naturally to the constrained-efficient case analyzed later.

We assume the planner maximizes a weighted average of expected lifetime utility across generations:

$$\mathcal{W}_0 = \sum_{t=0}^{\infty} \frac{1}{R^t} \sum_{i=1}^{J} \Lambda_i v_{it}^n (1-\omega) L_{it}, \tag{6}$$

where v_{it}^n denotes the lifetime value of the newborns born in location i at time t, $(1-\omega)L_{it}$ corresponds to the mass of households born in location i at time t, Λ_i is the welfare weight attached to households born in i, and $R > 1$ is the social discount rate.

Following the extensive literature on dynamic public finance, we employ a recursive formulation of the planning problem using promised utility and continuation value, instead of the sequence of allocations contingent on the history of preference shocks (e.g. Atkeson and Lucas 1992, Farhi and Werning 2007). At any given point in time t, each household is identified by their promised utility v and the location i. The planner assigns three sets of objects to each existing household. First, they determine the consumption level, $C_{it}(v)$. Second, they specify the migration destination plan, $\ell_{it+1}(v, \epsilon)$, which depends on the current promised utility v, the current location i, and the realization of preference shocks ϵ. Third, they choose the continuation utility for the next period, $v_{ijt+1}(v, \epsilon)$, again contingent on v, i, and ϵ. For newborns, the planner assigns analogous objects, except that they do not choose current consumption (as newborns do not consume in the birth period). Instead, they assign a lifetime utility.

Let $\phi_{it}(v)$ denote the measure of households in location i with promised utility v. The state variable of the planner is then $\phi_t \equiv [\phi_{it}(\cdot)]_i$. The planning problem in a recursive form follows

$$\mathcal{W}_t(\phi_t) = \max_{\{C_{it}(v),\ell^n_{it+1}(\epsilon),\ell_{it+1}(v,\epsilon),v^n_{it},v^n_{ijt+1}(\epsilon),v_{ijt+1}(v,\epsilon),l_{it},L_{it},\phi_{t+1}\}} \sum_i \Lambda_i v^n_{it}\,(1-\omega)L_{it} + \frac{1}{R}\mathcal{W}_{t+1}(\phi_{t+1}) \tag{7}$$

subject to

$$v^n_{it} = \beta\mathbb{E}_{it}\Big[\sum_j \mathbb{I}[\ell^n_{it+1}(\epsilon_{it}) = j]\left\{v^n_{ijt+1}(\epsilon_{it}) + \epsilon_{ijt}\right\}\Big] \tag{8}$$

$$v = u_{it}(C_{it}(v)) + \beta\omega\mathbb{E}_{it}\Big[\sum_j \mathbb{I}[\ell_{it+1}(v,\epsilon_{it}) = j]\left\{v_{ijt+1}(v,\epsilon_{it}) + \epsilon_{ijt}\right\}\Big] \tag{9}$$

$$\int C_{it}(v)d\phi_{it}(v) = f_{it}(\{l_{kit}, L_{kt}\}) \tag{10}$$

$$\sum_j l_{ijt} = \int d\phi_{it}(v) \tag{11}$$

$$L_{it} = \int d\phi_{it}(v) \tag{12}$$

and the law of motion of the distribution:

$$\begin{aligned}\phi_{jt+1}(v) = \sum_i \omega\mathbb{E}_{it}\left[\phi_{it}(v^{-1}_{ijt+1}(v,\epsilon_{it}))\mathbb{I}[\ell_{it+1}(v^{-1}_{ijt+1}(v,\epsilon_{it}),\epsilon_{it}) = j]\right] \\ + (1-\omega)L_{it}\mathbb{E}_{it}\left[\mathbb{I}[v^n_{ijt+1}(\epsilon_{it}) = v]\mathbb{I}[\ell^n_{it+1}(\epsilon_{it}) = j]\right],\end{aligned} \tag{13}$$

for all v. Constraints (8) and (9) correspond to the value functions of surviving households (1) and newborns (2), constituting the "promise-keeping constraints" for the planner. Note that the value of existing generations does not enter into the objective function because their values are predetermined when they are born. Constraints (10) and (11) are the resource constraint of the final goods and the labor services, respectively. Constraint (12) defines the population size.

While the original problem is quite complex due to the infinite dimensionality of the state and control variables, the problem dramatically simplifies when considering the Lagrangian of the above problem. We detail the derivations in Appendix A.1 and focus on the key results in the main text. Let P_{it}, w_{it}, and $w_{it}\alpha_{it}$ denote the Lagrange multipliers on (10), (11), and (12), respectively. Let $\mathcal{S}_t(\phi_t)$ denote the associated Lagrangian. We then guess and verify that the

Lagrangian $\mathcal{S}_t(\phi_t)$ is additively separable in (v, i):[7]

$$\mathcal{S}_t(\phi_t) = \sum_i \int S_{it}(v) d\phi_{it}(v) + D_t, \tag{14}$$

where D_t is the term that is independent of ϕ. The constant term D_t solves the following component problem:

$$D_t = \max_{\{l_{ijt}, L_{it}\}} \sum_i P_{it} f_{it}(\{l_{kit}, L_{kt}\}) - \sum_i w_{it} \sum_j l_{ijt} - \sum_i \alpha_{it} w_{it} L_{it} + \frac{1}{R} D_{t+1}. \tag{15}$$

Since D_{t+1} does not depend on $\{l_{ijt}, L_{it}\}$, this component problem is essentially a static problem. The first-order optimality conditions are given by the following static conditions:

$$P_{it} \frac{\partial f_{it}}{\partial l_{kit}} = w_{kt} \tag{16}$$

$$\sum_i P_{it} \frac{\partial f_{it}}{\partial L_{kt}} = w_{kt} \alpha_{kt}. \tag{17}$$

As one can imagine from the above expressions, and as we later discuss in Section 4.4 and Appendix B, the Lagrange multipliers correspond to prices in an equilibrium that decentralizes the planner's solution. That is, P_{it} corresponds to the price of final consumption goods in i, w_{it} corresponds to the wage in i, and α_{it} corresponds to the agglomeration elasticity in location i.

The dynamics come from the following component planning problem, which applies to each household in location i with promised utility v:

$$\begin{aligned} S_{it}(v) = \max_{C_{it}, \{\{v_{ijt+1}(\epsilon)\}, \ell_{it+1}(\epsilon)\}} & w_{it}(1 + \alpha_{it}) - P_{it} C_{it} + (1 - \omega) S^n_{it} \\ & + \frac{1}{R} \omega \mathbb{E}_{it} \Big[\sum_j \mathbb{I}[\ell_{it+1}(\epsilon_{it}) = j] S_{jt+1}\big(v_{ijt+1}(\epsilon_{it})\big) \Big] \end{aligned} \tag{18}$$

$$\text{s.t.} \qquad v = u_{it}(C_{it}) + \beta \omega \mathbb{E}_{it} \Big[\sum_j \mathbb{I}[\ell_{it+1}(\epsilon_{it}) = j] \left\{ v_{ijt+1}(\epsilon_{it}) + \epsilon_{ijt} \right\} \Big]. \tag{19}$$

The term S^n_{it} is the value function associated with newborns in location i, which solves

$$S^n_{it} = \max_{v^n_{it}, \{\{v^n_{ijt+1}(\epsilon)\}, \ell^n_{it+1}(\epsilon)\}} \Lambda_i v^n_{it} + \frac{1}{R} \mathbb{E}_{it} \Big[\sum_j \mathbb{I}[\ell^n_{it+1}(\epsilon_{it}) = j] S_{jt+1}(v^n_{ijt+1}(\epsilon_{it})) \Big] \tag{20}$$

$$\text{s.t.} \qquad v^n_{it} = \beta \mathbb{E}_{it} \Big[\sum_j \mathbb{I}[\ell^n_{it+1}(\epsilon_{it}) = j] \left\{ v^n_{ijt+1}(\epsilon_{it}) + \epsilon_{ijt} \right\} \Big]. \tag{21}$$

[7]Similar techniques appear in Atkeson and Lucas (1992), Farhi and Werning (2007, 2012), and Veracierto (2022, 2023) in the context of various different models.

The objective function $S_{it}(v)$ has a clear economic interpretation: it represents the "net social surplus" generated by a household in location i with promised utility v. Each such household contributes the marginal product of labor, w_{it}, and the agglomeration benefit, $w_{it}\alpha_{it}$, while incurring a resource cost equal to their consumption, $P_{it}C_{it}$. These terms can be interpreted as fiscal and technological externalities the household generates in the decentralized equilibrium that implements the planner's allocation. Additionally, households may die and be replaced by newborns, whose entry contributes to social surplus, as captured by the term $(1-\omega)S^n_{it}$. The continuation value reflects the possibility of migration across locations in future periods, making the expression recursive. The net social surplus of newborns S^n_{it} is defined similarly (equation (20)), but differences arise from the fact that (i) they do not work or consume in period t when they are born, and (ii) their lifetime value directly enters into net social surplus, as this is the first time they enter the economy.

We let $\{C_{it}(v), \{\{v_{ijt+1}(v,\epsilon)\}, \ell_{it+1}(v,\epsilon)\}\}$ denote the policy functions associated with (18). Since the promised utility v summarizes the relevant state of household h at each location i and time t, these policy functions, together with the realization of preference shocks $\epsilon_{it}(h)$, fully determine household h's consumption, continuation values (for each potential destination), and location choice: $\{C_{it}(h), \{v_{ijt+1}(h)\}, \ell_{it+1}(h)\}$. Taking first-order conditions and combining them with the envelope condition, we have the following characterization of the first-best allocation.

Proposition 1. *In the first-best, the following conditions must hold. For any household h living in location i at time t, consumption satisfies the following condition*

$$\frac{u'_{it}(C_{it}(h))}{P_{it}} = \beta R \frac{u'_{jt+1}(C_{jt+1}(h))}{P_{jt+1}} \tag{22}$$

for all j. The migration assignment solves

$$\max_{l} \left\{ \frac{P_{lt+1}}{u'_{lt+1}(C_{lt+1}(h))} \beta[v_{ilt+1}(h) + \epsilon_{ilt}(h)] + \frac{1}{R} S_{lt+1}(v_{ilt+1}(h)) \right\}. \tag{23}$$

The proof of this proposition, as well as those in the rest of the paper, appears in Appendix A. Proposition 1 reveals an intuitive property that the first-best allocation must feature. Condition (22) says that the marginal utility of resources must be equalized across space and over time after adjusting for discounting. If marginal utilities were not equalized, the planner could improve welfare by reallocating resources from locations or time periods where marginal utility is low to those where it is high. Because of the additive separability of idiosyncratic preference shocks, consumption does not depend on these shocks and is equalized within and across locations as well as over time.

The second condition (23) says that location choices must maximize social surplus. The first term captures the continuation value of households who move from i to l with idiosyncratic preference ϵ_{ilt}. It is multiplied by the inverse of the marginal utility of resources so as to denote the utils of continuation value in resource units. The second term captures the net social surplus of households in location l in period $t+1$ for the planner. In the first-best allocation, the planner allocates each individual to the location that maximizes the sum of these two.

4 Constrained-Efficient Allocations

The previous section assumed the planner perfectly observes each household's location preference shock and assigns their location accordingly. While it serves as a useful theoretical benchmark, such a situation is hardly realistic. We now turn to the optimal allocation when preference shocks are private information.

To see why private information matters, suppose that households can freely choose their location. Then, households make migration decisions to maximize their *private*, rather than *social*, surplus. Formally, the location choice now solves:

$$\max_{l} \left\{ v_{ilt+1} + \epsilon_{ilt} \right\}, \tag{24}$$

which clearly deviates from equation (23).

This implies that households would strategically misreport their preference shocks if they could. Households would pretend to have preference shocks such that they will be assigned to locations that solve (24) instead of (23). Consequently, the planner must respect incentive compatibility constraints in the presence of private information.

4.1 Incentive Compatibility Constraints

In this section, we first formalize the incentive compatibility constraints that the planner faces in this private information environment. By the revelation principle, we can focus on a direct revelation mechanism where households report their preference shock in each period. The incentive compatibility constraint requires that truth-telling is optimal from the household's perspective:

$$\epsilon \in \arg\max_{\hat{\epsilon}} \sum_{j} \mathbb{I}[\ell_{it+1}(v, \hat{\epsilon}) = j] \left\{ v_{ijt+1}(v, \hat{\epsilon}) + \epsilon_{ijt} \right\} \tag{25}$$

for each household in location i with promised utility v at time t. From this expression, it is tempting to think that the planner might want to promise a low continuation value in destination j, $v_{ijt+1}(v, \hat{\epsilon})$, to households reporting a high value of $\hat{\epsilon}_{ijt}$, since such households are likely to choose j regardless and thus require less incentive. However, equation (25) makes clear that this allocation violates incentive compatibility. In fact, $v_{ijt+1}(v, \hat{\epsilon})$ cannot depend on the reported preference shock, $\hat{\epsilon}$, conditional on migrating from i to j. If it did, households would simply (mis)report whichever preference gives the highest continuation value.

Therefore, any incentive compatible allocation must feature

$$v_{ijt+1}(v, \hat{\epsilon}) = v_{ijt+1}(v). \tag{26}$$

In other words, conditional on location choices and promised value, the planner cannot discriminate across households. Since the promised value v at each point in time is a function past location, the constrained-efficient allocation is achieved by policies that depend on the full history of location, but not directly on the history of preference shocks. As we explain in detail in Section 4.4, this result implies that the implementation of the constrained-efficient allocation need not involve preference reports. Instead, the implementation requires policies to be contingent only on the history of location choice.

Based on the above observation, we can rewrite the incentive compatibility constraint (25) as households directly choosing their location to maximize their utility based on the promised values, instead of reporting preferences and receiving location assignments:

$$\ell_{it+1}(v, \epsilon_{it}) \in \arg\max_{l} \ \{v_{ilt+1}(v) + \epsilon_{ilt}\} \tag{27}$$

for households living in location i with promised utility v and preference shocks ϵ_{it} at time t.

4.2 Recursive Formula for Constrained Efficiency

By imposing the incentive compatibility constraint (27) as an additional constraint in the problem (18) and noting that the next-period promised utility $v_{ijt+1}(v)$ cannot depend on preference shocks ϵ directly as discussed above, we can write the constrained-efficeint component planning problem as follows:

$$\begin{aligned} S_{it}(v) = \max_{C_{it}, \{v_{ijt+1}\}, \{\ell_{it+1}(\epsilon_{it})\}} & w_{it}\left(1+\alpha_{it}\right) - P_{it}C_{it} + (1-\omega)S_{it}^{n} \\ & + \frac{1}{R}\omega\mathbb{E}_{it}\left[\sum_j \mathbb{I}[\ell_{it+1}(\epsilon) = j] S_{jt+1}\big(v_{ijt+1}\big)\right] \end{aligned} \tag{28}$$

$$\text{s.t.} \qquad v = u_{it}(C_{it}) + \beta\omega\mathbb{E}_{it}\left[\sum_j \mathbb{I}[\ell_{it+1}(\epsilon_{it}) = j]\left\{v_{ijt+1} + \epsilon_{ijt}\right\}\right] \tag{29}$$

$$\ell_{it+1}(\epsilon_{it}) \in \arg\max_{l} \ \{v_{ilt+1} + \epsilon_{ilt}\}. \tag{30}$$

We can further simplify this problem using the representation result from Hofbauer and Sandholm (2002) and Donald et al. (2025). They show that any discrete choice problem can be equivalently represented as a maximization problem with respect to choice probability, subject to an appropriately defined cost function. Formally, we can rewrite the above problem as

$$S_{it}(v) = \max_{C_{it},\{v_{ijt+1},\mu_{ijt}\}} w_{it}\left(1+\alpha_{it}\right) - P_{it}C_{it} + (1-\omega)S_{it}^{n} + \frac{1}{R}\omega\sum_{j}\mu_{ijt}S_{jt+1}\left(v_{ijt+1}\right) \tag{31}$$

$$\text{s.t.} \qquad v = u_{it}(C_{it}) + \beta\omega\left[\sum_{j}\mu_{ijt}v_{ijt+1} - \psi_{it}(\{\mu_{ijt}\}_j)\right] \tag{32}$$

$$\{\mu_{ijt}\}_j \in \arg\max_{\{\tilde{\mu}_{ijt}\}_j}\left\{\sum_{j}\tilde{\mu}_{ijt}v_{ijt+1} - \psi_{it}(\{\tilde{\mu}_{ijt}\}_j)\right\}, \tag{33}$$

for some function $\psi_{it}(\cdot)$ that only depends on the distribution function of preference shocks $G_{it}(\cdot)$,[8] and $\mu_{ijt} \equiv \mathbb{E}_{it}\left[\mathbb{I}[\ell_{it+1}(\epsilon) = j]\right]$ is the probability that households in location i at time t to migrate to location j in the next period.[9]

This representation makes it clear that the problem has a mathematically similar structure to the optimal design of dynamic unemployment insurance, analyzed by Hopenhayn and Nicolini (1997) and Veracierto (2022). There, the planner seeks to equalize the marginal utility of the employed and unemployed over time, taking into account that doing so discourages job search effort. Here, the planner seeks to equalize marginal utility across space and over time, taking into account that doing so distorts migration choices. Despite different microfoundations for the incentive compatibility constraints, the planner faces a similar trade-off. An important difference is the dimensionality of the problem: households face many different options for migration, unlike single-dimensional job search effort. We revisit the computational challenges associated with this high-dimensionality in Section 6.

[8] See Hofbauer and Sandholm (2002) or Donald et al. (2025) for an explicit expression for $\psi_{it}(\cdot)$. For example, if $G_{it}(\cdot)$ is given by independent type-I extreme value distribution (logit), then $\psi_{it}(\{\mu_{ijt}\}_j) = \frac{1}{\theta}\sum_j \mu_{ijt}\ln(\mu_{ijt}/\chi_{ijt})$, where $\theta > 0$ is the scale parameter, and $\{\chi_{ijt}\}_j$ are location parameters capturing migration costs, as originally shown by Anderson, De Palma, and Thisse (1988).

[9] In general, $S_{it}(v)$ may not be concave in v, in which case, a lottery is needed to ensure concavity in the value function (see e.g. Prescott and Townsend 1984, Balke and Lamadon 2022). For the sake of notational simplicity, we abstract from the use of lotteries. In our quantitative exercise, we verify that $S_{it}(v)$ is concave, and therefore lotteries are not used even if available.

The net social surplus of newborns is represented similarly as

$$S_{it}^{n} = \max_{v_{it}^{n},\{v_{ijt+1}^{n},\mu_{ijt}^{n}\}_j} \Lambda_i v_{it}^{n} + \frac{1}{R}\sum_j \mu_{ijt}^{n} S_{jt+1}(v_{ijt+1}^{n}) \tag{34}$$

$$\text{s.t.} \qquad v_{it}^{n} = \beta \sum_j \left[\mu_{ijt}^{n} v_{ijt+1}^{n} - \psi_{it}(\{\mu_{ijt}^{n}\}_j)\right] \tag{35}$$

$$\{\mu_{ijt}^{n}\}_j \in \arg\max_{\{\tilde{\mu}_{ijt}^{n}\}} \left\{ \sum_j \left[\tilde{\mu}_{ijt}^{n} v_{ijt+1}^{n} - \psi_{it}(\{\tilde{\mu}_{ijt}^{n}\}_j)\right] \right\}, \tag{36}$$

where $\mu_{ijt}^{n} \equiv \mathbb{E}_{it}\big[\mathbb{I}[\ell_{it+1}^{n}(\epsilon) = j]\big]$. Again, the difference from the existing generation (31) arises from the fact that they do not work or consume in period t, and their lifetime value directly enters into net social surplus.

We let $\{C_{it}(v), \{v_{ijt+1}(v), \mu_{ijt}(v)\}\}$ denote the policy functions associated with the Bellman equation for the surviving generation (31). As in the first-best analysis of Section 3, these policy functions, together with the realization of preference shocks $\epsilon_{it}(h)$, fully determine household h's consumption $C_{it}(h)$ and continuation values (for each potential destination) $\{v_{ijt+1}(h)\}$. The following proposition provides a recursive formula that the constrained-efficient allocation must satisfy.

Proposition 2. *In any constrained-efficient allocation, the following must hold for all h, i, j, t:*

$$\mu_{ijt}\left[\beta R \frac{u'_{jt+1}(C_{jt+1}(h))/P_{jt+1}}{u'_{it}(C_{it}(h))/P_{it}} - 1\right] + \underbrace{\sum_k \frac{\partial \mu_{ikt}}{\partial C_{jt+1}} \frac{S_{kt+1}(v_{ikt+1}(h))}{P_{jt+1}}}_{\equiv\, \xi_{ijt}(h)} = 0, \tag{37}$$

where $\frac{\partial \mu_{ikt}}{\partial C_{jt+1}}$ is the derivative of migration from i to k with respect to consumption in j.

The above proposition characterizes the trade-off between equalizing marginal utility of income and distorting migration decisions. Without the term $\xi_{ijt}(h)$, the above condition is identical to the one for the first-best complete information benchmark (22) with full consumption smoothing. The term $\xi_{ijt}(h)$ captures an additional consideration the planner must address: how changes in consumption influence migration patterns and, consequently, net social surplus in the subsequent period. Specifically, increasing consumption at location j in period $t+1$ triggers migration responses that affect net social surplus. Importantly, these migration responses occur not only at location j, but across all other locations $k = 1, \ldots, J$ through substitution.

Consider first the case where $\xi_{ijt}(h) > 0$. This implies that raising consumption at location j in period $t+1$ induces migration responses that are beneficial in terms of net social surplus in the next period onward. In this case, the planner has an incentive to increase relative consumption at $(j, t+1)$ compared to (i, t), deviating from full consumption smoothing. That is, the planner back-loads consumption to encourage migration toward high-surplus locations. Conversely, when $\xi_{ijt}(h) < 0$, migration responses induced by higher consumption at $(j, t+1)$ reduce net social surplus. The planner then lowers relative consumption at $(j, t+1)$. That is, front-loading consumption to discourage migration to low-surplus locations.

To gain further intuition, consider a common special case where the preference shocks are drawn from an independent type-I extreme value (Gumbel) distribution with scale parameter $\theta > 0$, as in Artuç et al. (2010) and Caliendo et al. (2019). In this case, the expression for $\xi_{ijt}(h)$ simplifies to

$$\xi_{ijt}(h) = \theta \mu_{ijt} \frac{u'_{jt+1}(C_{jt+1}(h))}{P_{jt+1}} \left[S_{jt+1}(v_{ijt+1}(h)) - \sum_k \mu_{ikt} S_{kt+1}(v_{ikt+1}(h)) \right]. \tag{38}$$

Therefore, the consumption profile of this household if migrates to location j is back-loaded when the net social surplus of location j is higher than the migration probability weighted average net social surplus in all locations. Conversely, it is front-loaded when the net social surplus of j is lower than the weighted average. The extent of back-loading and front-loading naturally increases with higher migration elasticity θ.

Notice that our formula in Proposition 2 closely resembles the optimal unemployment insurance formula of Baily (1978) and Chetty (2006). In their setting, increasing unemployment insurance discourages the incentive to transition out of unemployment, creating a loss in net social surplus (i.e. a fiscal externality). Optimal unemployment insurance thus balances the benefits of consumption smoothing against the incentive to transition back to employment. Our formula extends this logic to the context of dynamic discrete choice. Similarly, our expression relates to formulas derived in the literature on optimal wage-tenure contracts between firms and workers (Burdett and Coles 2003, Balke and Lamadon 2022, and Souchier 2022). In those settings, a similar trade-off arises between providing insurance (via wage smoothing) and preserving incentives to retain or separate workers. Finally, our formula serves as a dynamic counterpart to the static optimal spatial transfer policies studied by Ales and Sleet (2022) and Donald et al. (2025), where the planner aims to equalize marginal utility across space.

Together with the formula in Proposition 2, the rest of the allocation and the Lagrange multipliers are determined in the same way as the complete information first-best benchmark. In particular, $\{l_{ijt}, L_{it}, \phi_{it}, P_{it}, w_{it}, \alpha_{it}\}$ solve (10), (11), (12), (13), (16), and (17).

To further understand the property of the constrained-efficient allocation, it is useful to observe that Proposition 2 also implies the following property of for consumption.

Corollary 1. *In any constrained-efficient allocation, the following equation must hold for each* *household h living in location i at time t:*

$$\frac{P_{it}}{u'_{it}(C_{it}(h))} = \mathbb{E}_{it}\left[\frac{P_{jt+1}}{\beta R u'_{jt+1}(C_{jt+1}(h))}\right]. \tag{39}$$

This equation is commonly referred to as the "inverse Euler equation", as it equates the inverse of the marginal utility of resources across time (Diamond and Mirrlees 1978, Rogerson 1985, Golosov, Kocherlakota, and Tsyvinski 2003, Farhi and Werning 2012, Bloedel et al. 2025). It is a common feature of constrained-efficient allocations in economies with private information and additively separable preferences.

The inverse Euler equation is also useful for interpreting the distortions, or "wedges", that arise in decentralized market economies. In an economy with a risk-free asset (as studied in Section 6), households' consumption satisfies the standard *Euler equation*, not the *inverse* Euler equation, which generally do not coincide. As emphasized by Golosov et al. (2003) and Farhi and Werning (2012), the inverse Euler equations typically imply more front-loaded consumption than the decentralized equilibrium, on average.

4.3 History-Independent Constrained-Efficient Allocation

Proposition 2 shows that constrained-efficient allocations generally require consumption to depend on a household's migration history, summarized by its promised utility and current locations. We now consider an alternative allocation, which we refer to as the "history-independent constrained-efficient allocation", in which consumption depends only on a household's current location, not its full migration history. Although this allocation is obviously suboptimal relative to the fully history-dependent case, analyzing it is useful for at least three reasons. First, comparing it to the fully history-dependent case highlights the importance of the dynamic incentives emphasized earlier. Second, this case illustrates how our approach can be extended to accommodate further restrictions of the policy space. Third, such a policy might be attractive for practical reasons, as policymakers may face informational or administrative constraints that limit their ability to condition policies on households' migration histories.

Specifically, we impose the restriction that all households residing in the same location must receive the same level of consumption, regardless of their migration history; that is, $C_{it}(v) = C_{it}$. This assumption implies that all households currently living in location i attain the same expected utility going forward. As a result, the component planning problems described in Section 4.2 become degenerate for v, and the net social surplus within each location is likewise uniform, i.e. $S_{it}(v) = S_{it}$. Aside from this restriction, the rest of the problem remains unchanged. We formally define the planning problem and derive the optimality conditions in Appendix A.5. This results in the following condition that any history-independent optimal policy must satisfy.

Proposition 3. *In any history-independent constrained-efficient allocation, the path of consumption in each location satisfies*

$$\sum_i L_{it}\left[\mu_{ijt}\left[\beta R\omega\frac{u'_{jt+1}(C_{jt+1})/P_{jt+1}}{u'_{it+1}(C_{it})/P_{it}}+\beta R(1-\omega)\Lambda_i\frac{u'_{jt}(C_{jt+1})}{P_{jt+1}}-1\right]+\sum_k\frac{\partial\mu_{ikt}}{\partial C_{j+1}}\frac{S_{kt+1}}{P_{jt+1}}\right]=0, \tag{40}$$

where

$$S_{jt}=w_{jt}\left(1+\alpha_{jt}\right)-P_{jt}C_{jt}+(1-\omega)\Lambda_j v^n_{jt}+\frac{1}{R}\sum_k\mu_{jkt}S_{kt+1}. \tag{41}$$

The formula in Proposition 3 resembles that in Proposition 2 in that it characterizes the deviation from full consumption smoothing. The two important differences are (i) it is averaged across migration origin i, since next-period consumption C_{jt+1} cannot depend on where the household migrated from, and (ii) it is implicitly averaged across households with any migration history, since C_{it}, C_{jt+1}, and S_{kt+1} cannot depend on v. These restrictions imply that the planner pools the insurance-incentive trade-offs across all households currently in the same location.

As a result, the planner's ability to front-load or back-load consumption for individual households is limited. For example, if the planner wants to attract more households to a particular location, they must increase consumption in that location uniformly, regardless of each household's prior location. In Section 6, we quantify how these constraints shape the allocation of consumption and population in our application to the US state-level economy.

Our approach can naturally be extended to accommodate alternative restrictions on the policy space. In Appendix C, we study an intermediate case in which the planner specifies consumption by origin, destination, and time, C_{ijt}. This lies between the fully history-dependent constrained-efficient allocation in Section 4.2 and the history-independent allocation in Section 4.3. We derive the analogous formula for this setting, which characterizes the trade-off between consumption smoothing and efficient migration. Compared to the history-independent case, there is more scope for dynamic incentives, as consumption can vary by origin. Nonetheless, when applied to the US economy, we find that the resulting population allocation is nearly identical to the history-independent case, suggesting that escaping the insurance-incentive trade-off requires richer dynamic policies.

4.4 Implementation

Thus far, we have focused on characterizing the constrained-efficient allocations by taking incentive compatibility in migration decisions as a fundamental constraint, without specifying the underlying market structure or available policy instruments. This approach highlights the core policy trade-off and establishes a benchmark for the best outcome a planner can achieve, independent of implementation details. In this section, we briefly discuss how such constrained-efficient allocations might be implemented under different market structures.

We divide the discussion into two types of environment, based on the existing dynamic spatial equilibrium literature. First, researchers have extensively considered environments where households making dynamic migration decisions are hand-to-mouth (e.g. Caliendo et al. 2019). Second, and more recently, researchers have started to consider environments where agents make both dynamic migration and consumption-saving decisions simultaneously (e.g. Giannone et al. 2023, Dvorkin 2023, Greaney 2023, Greaney et al. 2025).

In the first environment, the constrained-efficient allocation can be most naturally implemented through direct transfers to households. As we formally show in Appendix B, these transfers need to be conditioned on households' location histories. Similarly, the history-independent constrained-efficient allocation can be implemented with transfers that depend only on a household's current location and time. Thus, simple place-based transfers are sufficient to implement the history-independent constrained-efficient outcome.

In the second environment, where agents make both migration and consumption-saving decisions, transfers alone generally cannot implement either the constrained-efficient or the history-independent constrained efficient allocation. This is because changes in transfers affect not only migration choices but also consumption and saving behavior. Such joint deviations cannot be controlled with a single policy instrument such as transfers, so additional tools are required (Stantcheva 2020). For instance, a capital income tax that depends on the history of residential locations can implement the constrained-efficient allocation, as shown by Kocherlakota (2005) in the context of a dynamic Mirrleesian framework. In our quantitative analysis in Section 6, we sidestep this implementation question and focus on comparing the constrained-efficient allocation to the status quo allocation.[10]

In both environments, an important aspect of implementation is the history dependence in policy design. While a complete implementation of history dependence may run up against informational and administrative constraints, simplified forms of history-dependent policies are commonly observed. For example, advanced economies such as Japan and parts of Europe offer

[10] As Golosov et al. (2003) note, "the robust predictions of any kind of theory of optimal taxation are not about *taxes*, but, like our results, are *wedges*".

migration subsidies tied to minimum residency periods in declining or depopulated regions.[11] Similarly, access to urban public services is sometimes restricted only to long-term residents, as in China's Hukou system. Our analysis of constrained-efficient allocations characterizes the best outcomes achievable by this class of policies.

5 Extensions

Our baseline model deliberately abstracts from several considerations in order to transparently convey the main trade-off in the constrained-efficient allocation. We now discuss extensions and generalizations of the baseline model.

5.1 Capital Accumulation

Some existing work (e.g. Kleinman et al. 2023, Bilal and Rossi-Hansberg 2023, D'Amico and Alekseev 2024) introduce capital accumulation into the dynamic spatial equilibrium environment, which we have abstracted from so far. In Appendix D.1, we show that our environment can be straightforwardly extended to incorporate location-specific capital accumulation (e.g. building structure or housing stocks) subject to adjustment costs. Importantly, such considerations do not meaningfully interact with the trade-off that we highlighted in Proposition 2. In fact, Proposition 2 and the underlying Bellman equations remain unchanged. Meanwhile, optimal investment and capital accumulation follow the standard q-theory of investment.

5.2 Ex-ante Heterogeneous Households

In the baseline model, we have assumed that all households are ex-ante homogeneous. In Appendix D.2, we extend our baseline environment to an environment with many ex-ante heterogeneous household types $\theta \in \{\theta_1, \dots, \theta_M\}$ with arbitrary heterogeneity in preferences, location choice, and productivity. We also consider a general form of agglomeration/congestion forces that allow for spillover across different household types. There, we show that Proposition 2 and the underlying Bellman equations remain unchanged, except that now everything is indexed by θ.

5.3 Lagged Agglomeration/Congestion Forces

In the baseline model, we have assumed that the agglomeration and congestion forces arise from contemporaneous population size. Some existing work, such as Allen and Donaldson (2020) and

[11] For example, the Japanese government offers one million yen (approximately 10,000 USD) to households that migrate out of Tokyo and commit to stay in the designated depopulated regions for at least five years (https://www.chisou.go.jp/sousei/ijyu_shienkin.html (in Japanese)).

Peters (2022), allows for agglomeration/congestion forces that depend on the lagged population size. In Appendix D.3, we extend our environment to allow for agglomeration/congestion forces that depend on arbitrarily long lags of the population size distribution. The only material difference from our baseline model is that the net social surplus from agglomeration/congestion externalities becomes the discounted sum of all future externalities, not only the contemporaneous ones $w_{it}\alpha_{it}$.

6 Quantification

In this section, we calibrate our model to the US state level and quantitatively assess how the constrained-efficient allocation differ from that of the status quo economy.

6.1 Status Quo Economy with Migration and Savings

We assume that the data are generated from a competitive equilibrium, in which households make migration decisions and consumption-saving decisions, which we call the "status quo economy". First, households make forward-looking migration decisions to maximize utility, as in Caliendo et al. (2019). Second, households choose consumption and savings with risk-free assets and occasionally binding borrowing constraints, following the (exogenously) incomplete market literature (Bewley 1986, Huggett 1993, Aiyagari 1994, Imrohoroğlu 1989).

The only available assets in the economy are state non-contingent bonds in zero net supply. Let $1 + r_t$ be the gross rate of return from bond holdings from time t to $t + 1$. All households face a common exogenous borrowing limit, with the minimum asset level given by $\underline{a}$. In addition, the government provides spatial transfers: households living in location j at time t receive transfers T_{jt}, independent of their asset holdings.

The Bellman equation of households living in location j at time t with asset holdings a_t is

$$v_{it}(a_t) = \max_{C_{it},\{\mu_{ijt}\},a_{t+1}} u_{it}(C_{it}) + \beta\omega\left[\sum_j \mu_{ijt}v_{jt+1}(a_{t+1}) - \psi_{it}(\{\mu_{ijt}\})\right] \tag{42}$$

$$\text{s.t.} \quad P_{it}C_{it} + a_{t+1} = (1 + r_{t-1})a_t + w_{it} + T_{it} \tag{43}$$

$$a_{t+1} \geq \underline{a}, \tag{44}$$

where P_{it} is the price index and w_{it} is the wage. Here, we adopt the same representation of the dynamic discrete choice problem as in Section 4.2, where households directly choose migration probabilities $\{\mu_{ijt}\}_j$ subject to the cost function ψ_{it}. Let $C_{it}(a)$, $a_{it+1}(a)$, and $\mu_{ijt}(a)$ denote the policy functions associated with the above value functions. We assume that when households die, the asset is transferred as accidental bequests to their offspring.

Since we assume that assets are in zero net supply, our model nests commonly used hand-to-mouth households as a special case with $\underline{a} = 0$, under which no household can borrow or save in equilibrium. The traditional Bewley-Hugget-Aiyagari model is nested as a special case when the migration is completely inelastic.

There is a representative firm in each location that imports factor services from other regions and produces non-traded final goods. The representative firm in location j solves

$$\max_{\{l_{kjt}\}_k} P_{jt} f_{jt}(\{l_{kjt}\}_k, \{L_{kt}\}_k) - \sum_k w_{kt} l_{kjt}, \tag{45}$$

taking the population sizes $\{L_{kt}\}_k$ as given. Here, agglomeration/congestion forces are externalities that are not internalized by private agents.

We assume that the government runs a balanced budget. The government budget constraint is

$$\sum_i T_{it} L_{it} = 0. \tag{46}$$

Let $\varphi_{jt}(a)$ denote the measure of households with asset level $a \in \mathbb{A}$ in location j. The goods market clearing condition is

$$\int C_{jt}(a) d\varphi_{jt}(a) = f_{jt}(\{l_{kjt}, L_{kt}\}). \tag{47}$$

The consistency of population size requires

$$L_{kt} = \int d\varphi_{kt}(a). \tag{48}$$

The factor market clearing condition is

$$\sum_j l_{kjt} = L_{kt}. \tag{49}$$

The distribution evolves according to the following law of motion:

$$\varphi_{jt+1}(a) = \sum_i \mu_{ijt}(a^{-1}_{jt+1}(a)) \varphi_{it}(a^{-1}_{jt+1}(a)). \tag{50}$$

The decentralized equilibrium of the status quo economy consists of value and policy functions $\{v_{it}(a), C_{it}(a), a_{it+1}(a), \mu_{ijt}(a)\}$, factor contents of trade $\{l_{kjt}\}$, population distributions

$\{L_{kt}\}_k$, spatial transfers $\{T_{it}\}$, distribution over assets in each location $\{\varphi_{jt}\}$, and prices $\{w_{it}, P_{it}, r_t\}$ such that: (i) given prices $\{w_{it}, P_{it}, r_t\}$ and policy $\{T_{it}\}$, the value and policy functions $\{v_{it}(a), C_{it}(a), a_{it+1}(a), \mu_{ijt}(a)\}$ solve the household's problem (42); (ii) given prices $\{w_{it}, P_{it}\}$ and population size $\{L_{kt}\}$, the factor contents of trade $\{l_{kjt}\}$ solve the firm's problem (45); (iii) transfers satisfy the government's budget constraint (46); (iv) markets clear (47), (49); (v) the population size is consistent (48); and (vi) the distribution $\{\varphi_{jt}\}$ evolves according to (50).

The status quo economy generally does not achieve either the first-best or constrained-efficient allocation, for two main reasons. First, private agents do not internalize agglomeration or congestion externalities when making migration decisions. Second, markets are incomplete: there is no mechanism to insure agents against idiosyncratic preference shocks and the resulting uncertainty in location choice. In what follows, we quantify these deviations through the lens of the calibrated model.

6.2 Calibration

We calibrate our status quo economy to match US data in 2017, assuming that the US is in steady-state. One period is five years. We consider the 48 contiguous states as a geographical unit, excluding the states of Alaska and Hawaii. We choose Alabama's labor as numeraire and set its wage to one. Since our calibration assumes a steady-state, we drop the subscript t in this subsection. We reintroduce the time subscripts when we study transition dynamics in Section 7.

We first parameterize the utility function as a CRRA utility function,

$$u_j(C_j) = \frac{C_j^{1-\gamma} - 1}{1-\gamma}, \tag{51}$$

and set $\gamma = 1$, corresponding to log utility, a standard specification in the literature (e.g. Caliendo et al. 2019). We also set ω so that the average life expectancy is 70 years. The production function is assumed to take the constant elasticity of substitution form:

$$f_j(\{l_{kj}\}, \{L_k\}_k) = \left[\sum_k \left(\mathcal{A}_{kj}(L_k)\, l_{kj}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{52}$$

where $\sigma > 1$ corresponds to the trade elasticity, and $\mathcal{A}_{kj}(L_k)$ is the productivity shifter of goods shipped from location k to j that depends on the population size of location k. This productivity is an iso-elastic function of the population size in the origin location:

$$\mathcal{A}_{kj}(L_k) = A_{kj} L_k^{\alpha}, \tag{53}$$

where A_{kj} is the fundamental component of productivity and can include iceberg trade costs. This specification, together with (16) and (17), implies

$$\alpha_j = \alpha \quad \text{for all } j. \tag{54}$$

For the baseline model, we assume $\alpha = 0.02$. This value corresponds to the lower end of the estimates in the literature summarized in Melo, Graham, and Noland (2009). We choose the lower end of the existing estimates as a baseline because, given our analysis is at the state-level, agglomeration forces are likely to be weaker than what is implied by the typical estimates at the city-level. We set the trade elasticity to $\sigma = 5$, as in Costinot and Rodríguez-Clare (2014).

Following Artuç et al. (2010) and Caliendo et al. (2019), we assume $\{\epsilon_j\}_j$ follow an independent type-I extreme value distribution, which implies the following migration cost function:

$$\psi_i(\{\mu_{ij}\}_j) = \frac{1}{\theta} \sum_j \mu_{ij} \ln(\mu_{ij}/\chi_{ij}), \tag{55}$$

as originally shown by Anderson et al. (1988). The parameter θ governs the migration elasticity, and $\chi_{ij} > 0$ represents the bilateral migration cost shifter. We set the value of migration elasticity to $\theta = 2.5$ for a five-year horizon, in line with the parameters estimated and used in the previous literature (Caliendo et al. 2019, Kleinman et al. 2023).

We briefly describe the calibration of the other parameter values and relegate the details to Appendix F.2. We choose $\{A_{ij}\}_{i \neq j}$ and $\{\chi_{ij}\}_{i \neq j}$ to match bilateral trade and migration flows at the state level. We choose $\{A_{ii}\}$ to match the real wage level in each state, and we normalize $\chi_{ii} = 1$ for all i. We set the discount factor β so that the annual real interest rate is 3%. We choose the value of $\underline{a}$ so that the average marginal propensity to consume (MPC) over a five-year horizon is 0.7, which lies in the middle of estimates in the literature.[12] We parameterize the transfer as $T_j = \varkappa_j w_j + \bar{T}$ and choose $\{\varkappa_j\}$ to match the ratio of net transfers to income at the state level. We obtain net transfers from the government at the state level from the Bureau of Economic Analysis (BEA). We adjust the term $\bar{T}$ to ensure that the government budget constraint holds.

Finally, we need to take a stance on the parameters that govern the social welfare function, $\{R, \Lambda_j\}$. We choose the social discount factor $1/R$ to be the same as the private discount factor β, so that $\beta R = 1$. As one can see from Proposition 1 and explored by Eden (2023) in depth, the first-best allocation features an equitable steady-state distribution of consumption across age

[12]To the best of our knowledge, Fagereng, Holm, and Natvik (2021) is the only study that estimates MPC over a five-year horizon. They document that households spend roughly 90% on lottery winnings over five years following the winning. Other studies (e.g, Orchard, Ramey, and Wieland 2023, Colarieti, Mei, and Stantcheva 2024, Boehm, Fize, and Jaravel 2025) find MPCs are 30-50% over a quarter or a year with no further statistically significant spending responses over longer (but less than five years) horizons.

Table 1: Parameter Values

Parameter	Description	Value	Source/Target
A. Assigned Parameters			
γ	Risk aversion	1	Standard
σ	Trade elasticity	5	Costinot and Rodríguez-Clare (2014)
θ	Migration elasticity	2.5	Caliendo et al. (2019)
ω	Surviving probability	0.93	Life expectancy 70 years
α	Agglomeration elasticity	0.02	Baseline
B. Internally Calibrated Parameters			
β	Private discount factor	0.92	Annual real interest rate 3%
$\underline{a}$	Borrowing limit	-0.08	5-year MPC 0.7
$\{A_{ij}\}$	Productivity shifter	-	Trade flows and real output
$\{\chi_{ij}\}$	Migration cost shifter	-	Migration flows
$\{\varkappa_i\}$	Net transfer rate	-	Net transfer from the government
C. Parameters for Social Welfare Function			
$1/R$	Social discount factor	0.92	Private discount factor
$\{\Lambda_j\}$	Location welfare weights	-	Equal weight

Notes: This table shows the parameters used in our quantitative exercise. Parameter values for $\{A_{ij}\}$, $\{\chi_{ij}\}$, and $\{\varkappa_i\}$ are chosen to exactly match the data moments described in the main text. One period corresponds to five years.

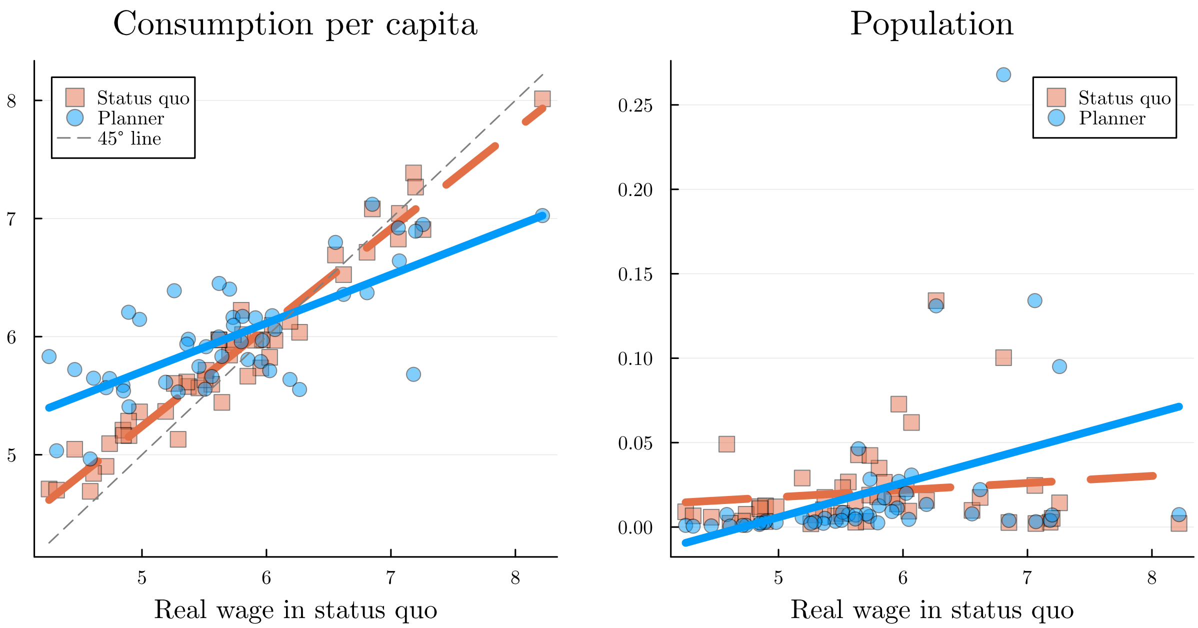
groups if and only if $\beta R = 1$. By setting $\beta R = 1$, any non-constant consumption-age profile in the constrained-efficient allocation can be attributed solely to private information. The planner puts equal weight on households born in different regions, which we normalize to one: $\Lambda_j = 1$ for all j.

6.3 Steady-State Computational Algorithms

Here, we briefly describe how we compute the steady states (for both the status quo and constrained-efficient allocation) and relegate the details to Appendix F.3. We first guess wages $\{w_j\}$ and population size $\{L_j\}$ by location. This allows us to obtain price indices $\{P_j\}$. We then iteratively solve the Bellman equations, (31) and (42), to obtain policy functions. With policy functions in hand, we can compute the steady-state distributions. We then check the market clearing conditions and consistency of the population size distribution. We then update the guess of $\{w_j\}$ and $\{L_j\}$ and repeat the whole procedure until both conditions are satisfied.

A key computational challenge lies in solving the Bellman equations. The Bellman equation for a constrained-efficient allocation (31) involves maximizing over $2J + 1$ control variables $(C_{it}, \{v_{ijt+1}, \mu_{ijt}\})$. The Bellman equation for the status quo economy (equation (42)) has fewer

Figure 1: Steady-State Consumption and Population: Status Quo vs. Planner



Notes: The left panel plots the average consumption per capita in each state against the real wage in the status quo economy. The square dots correspond to the status quo economy, and the circular dots correspond to the planner's solution. The dashed red line is the best linear fit for the status quo economy, and the solid blue line is the best linear fit for the planner's solution. The right panel plots population size against the real wage in the status quo economy and is analogous to the left panel.

dimensions, involving $J + 2$ control variables $(C_{it}, \{\mu_{ijt}\}, a_{t+1})$, but still poses a similar computational burden. Since we need to repeatedly solve the Bellman equations to find the fixed point in wages $\{w_j\}$ and the population distribution $\{L_j\}$, naive algorithms such as grid search or Newton's method are infeasible in our context.

To address this challenge, we extend the endogenous gridpoint method by Carroll (2006), which exploits the analytical first-order conditions to avoid root-finding or explicit optimization. For the status quo economy, our formulation of dynamic discrete choice as a direct optimization over migration probabilities $\{\mu_{ij}\}$ allows for a seamless extension.[13] For the constrained-efficient allocation, the extension is nontrivial due to the added dimensionality from destination-specific promised utilities $\{v_{ijt+1}\}$. Nevertheless, we demonstrate that the method can be tractably extended to the planner's problem under type-I extreme value preference shocks. This achieves substantial speed gains over naive algorithms, making it feasible to compute the steady-state allocation.

6.4 Steady-State: Status Quo vs. Planner

Armed with our calibrated parameters, we begin by comparing the steady-state of the status quo economy with that of the constrained-efficient allocation. The left panel of Figure 1 plots average consumption per capita against real wages in the status quo economy, which serve as a proxy for how "productive" each location is.[14] Square dots show average consumption in each state under the status quo. If the economy were in financial autarky, these points would lie on the 45-degree line. Instead, access to savings and government transfers smooth consumption, resulting in a slope slightly below one. Circular dots depict consumption in the constrained-efficient allocation, whose gradient is significantly flatter. This pattern indicates that the constrained-efficient allocation achieves substantially greater spatial consumption equality.

The right panel of Figure 1 compares population sizes, plotted against real wages in the status quo economy. In the status quo, the relationship is roughly flat. By contrast, the constrained-efficient allocation exhibits an upward-sloping relationship: more productive states host larger populations, and less productive states host smaller ones.

At first glance, this pattern may seem counterintuitive. If one has a static view of the world, where households base their location decisions solely on average consumption, one would expect the opposite pattern. *Lower* average consumption should mean *lower* population in productive states. In contrast, Figure 1 shows that the constrained-efficient allocation results in *lower* average consumption per capita yet *higher* population in productive states. How can the constrained-efficient allocation both reduce spatial consumption inequality *and* allocate more of the population to high productivity locations?

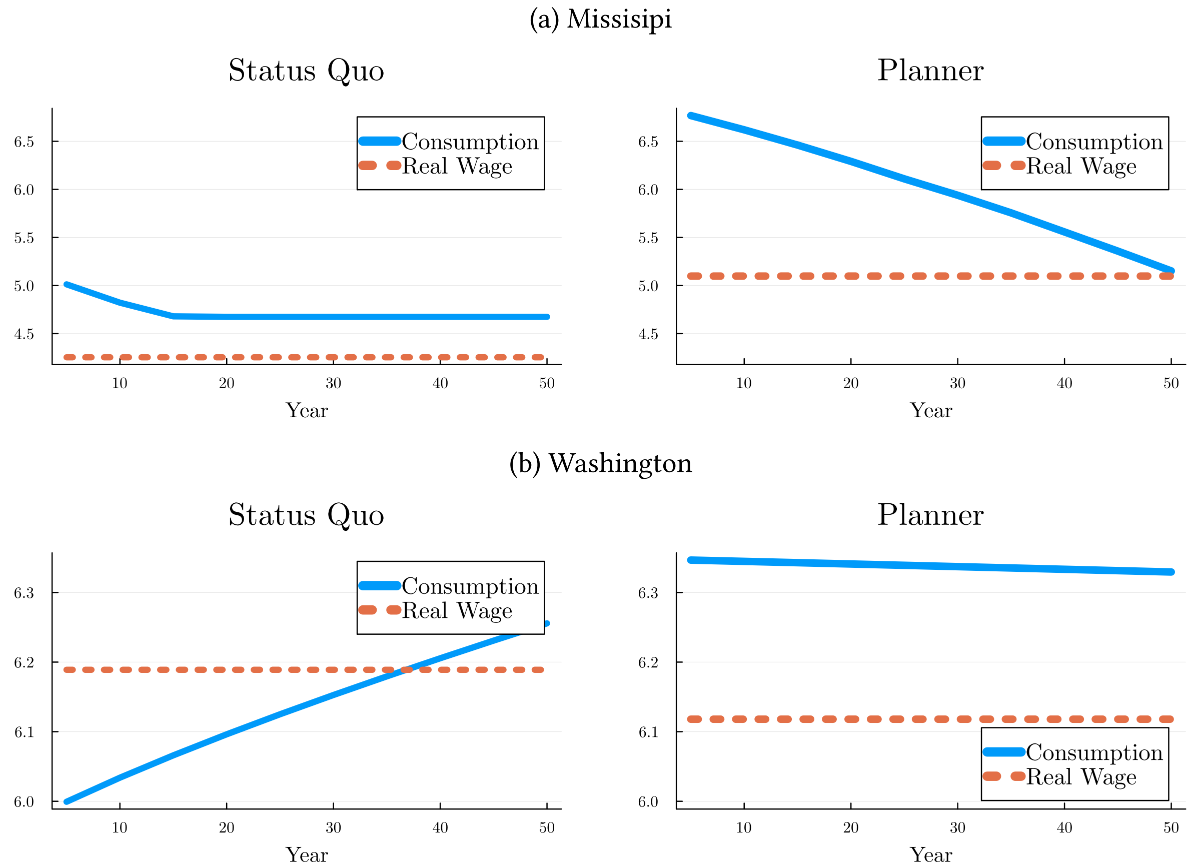
The answer lies in the role of dynamic incentives. As emphasized in Proposition 2, the planner front-loads consumption in unproductive states. This threat to cut consumption in the future incentivizes households to out-migrate without cutting consumption today. These incentives support greater reallocation toward productive regions while simultaneously improving consumption insurance in less productive areas. In doing so, dynamic incentives achieve both spatial efficiency and equity. This pattern marks a sharp contrast with static settings, where these goals are typically in conflict (Gaubert et al. 2021, Ales and Sleet 2022, Donald et al. 2025).

To further illustrate the role of dynamic incentives, Figure 2 plots life-cycle consumption and income profiles in Mississippi (an "unproductive" location) and Washington (a "productive" location). We consider households born in these locations who end up staying there for their lifetime. In the status quo economy, households born in Mississippi borrow to consume more

[13] See Greaney (2023) and Greaney et al. (2025) for a related continuous-time implementation in similar market environments.

[14] Appendix Figure F.2 confirms a strong positive relationship between real wages and average net social surplus S_j for the constrained-efficient allocation. Appendix Figure F.1 shows that the shadow value of labor w_j in the planner's solution closely tracks those in the status quo.

Figure 2: Consumption and Real Wages over the Life-Cycle for Stayers



Notes: The figure plots the consumption and real income profiles of households born in Mississippi (panel (a)) and Washington (panel (b)). We focus on the households who remain in the same location. The left panel shows the status quo economy, and the right panel shows the planner's solution (constrained-efficient allocation). In both cases, the initial value is the average of the households born in each location. The real wage in the planner's solution corresponds to the ratio of the Lagrange multipliers for labor and consumption goods (w_i/P_i).

than their real wages. This immediately drives down their assets which eventually hit the borrowing constraint. After hitting the borrowing constraint, consumption remains flat. In contrast, in the planner's solution, the consumption profile is front-loaded with an initially higher level of consumption than the status quo economy. In this way, the planner can effectively insure households born in Mississippi while strongly incentivizing them to leave Mississippi. The Washington example illustrates the other case. In the status quo economy, consumption increases over the life-cycle as households accumulate savings. The planner solution features nearly a flat consumption profile, or a far less front-loaded consumption profile than Mississippi, which incentivizes households to stay in Washington.

We quantify the aggregate consequences of moving to the constrained-efficient allocation in Table 2. The constrained-efficient allocation features around 9% higher real GDP at the national

Table 2: Steady-State Aggregate Outcomes

	Real GDP		$\text{Var}_L(\ln \bar{C}_j)$	
	Level	Change	Level	Change
0. Status quo	5.86	-	0.0094	-
1. Planner	6.38	+8.9%	0.0065	-31.3%
2. History independent	6.36	+8.6%	0.0140	+49.3%

Notes: This table shows aggregate real GDP and spatial inequality, measured by the variance of log average state-level consumption ($\bar{C}_j$). We show them for the status quo economy, (history-dependent) constrained-efficient allocation, which we label "planner", and the history-independent constrained-efficient allocation. We compute real GDP as nominal GDP divided by the weighted average of state-level price indices, where the weight is state-level nominal GDP. The variance of log state-level consumption is weighted by state population. Formally, $\text{Var}_L(\bar{C}_j) \equiv \sum_{j=1}^{J} L_j(\ln \bar{C}_j - \mathbb{E}_L[\ln \bar{C}_j])^2$ with $\mathbb{E}_L[\ln \bar{C}_j] \equiv \sum_{j=1}^{J} L_j \ln \bar{C}_j$.

level than the status quo economy. As emphasized earlier, this does not come at the cost of increasing spatial inequality. In fact, the constrained-efficient allocation cuts the variance of log regional consumption by a third. In terms of welfare, we find that utilitarian welfare of newborns increases by 9.5% in consumption equivalent units.[15] We further decompose these welfare gains into within- and between-location components in Appendix F.6, which we find to be equally important for welfare gains.

6.5 History-Independent Constrained-Efficient Allocation

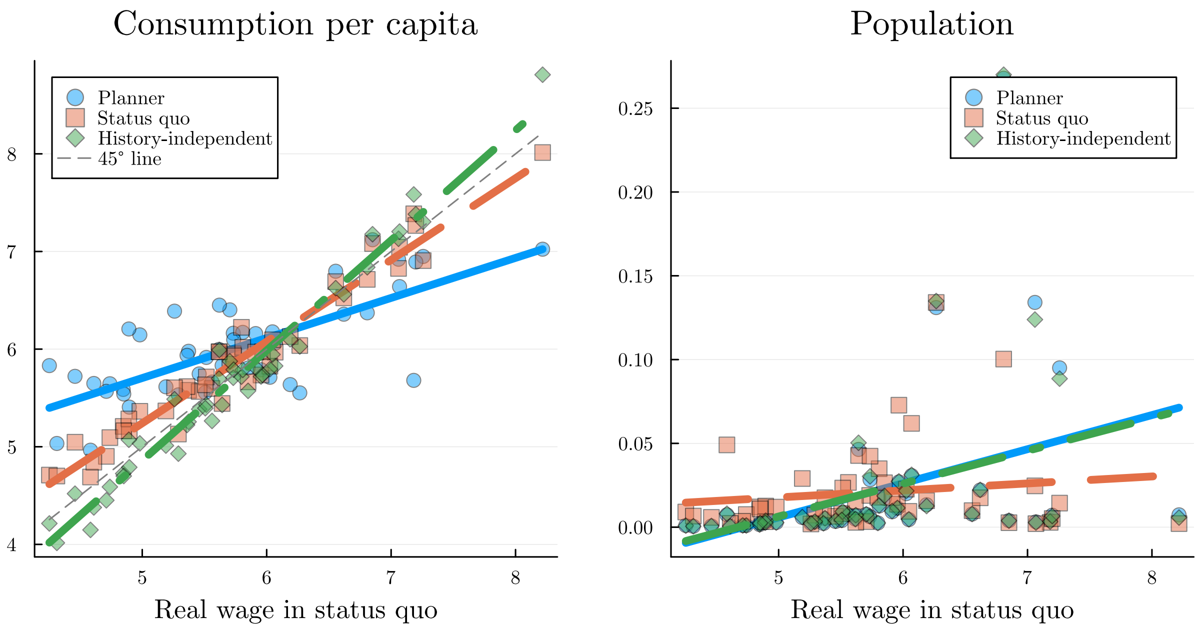
We next compare the history-contingent constrained-efficient allocation with the history-independent allocation, in which consumption depends solely on the household's current location, as described in Section 4.3. Without the ability to condition on migration history, the planner is constrained in front-loading or back-loading consumption. This comparison highlights the central role of dynamic incentives in achieving efficient outcomes.

Figure 3 presents per capita consumption (left panel) and population size (right panel) under the history-independent constrained-efficient allocation, in a format analogous to Figure 1. For comparison, we also include the status quo and the history-dependent constrained-efficient (planner) allocation.

The left panel shows that the history-independent allocation generates substantial spatial inequality in consumption, with a slope steeper than the 45-degree line. In this allocation, per capita consumption is *higher* in more productive states and *lower* in less productive ones, relative to both the status quo and the history-dependent constrained-efficient allocation. Despite

[15]This welfare number should be interpreted with caution, as it does not take into account the costs incurred during the transition, which may have been borne by previous generations.

Figure 3: Steady-State Consumption and Population: Planner vs. History-Independent



Notes: The left panel plots the average consumption per capita in each state against the real wage in the status quo economy. The circular, square, and diamond dots correspond to the history-dependent, status quo, and history-independent allocations, respectively. The right panel plots the population and is analogous to the left panel.

this greater inequality, the right panel demonstrates that the history-independent allocation still results in larger populations in more productive states and smaller populations in less productive ones. In fact, the population distribution is remarkably similar to that of the history-dependent allocation, even though their consumption patterns differ sharply.

These contrasts highlight the role of dynamic incentives. Without the ability to condition on migration history, the planner faces a sharper trade-off between spatial equity and migration efficiency. In our calibration, to reap efficiency gains, the planner sacrifices consumption smoothing in less productive regions. By contrast, the history-dependent allocation leverages dynamic incentives to achieve both spatial efficiency and equity.[16]

The last row of Table 2 quantifies the aggregate consequences of history-independent policy. Real GDP increases slightly less than in the constrained-efficient allocation. At the same time, spatial inequality is nearly 50% higher than in the status quo economy. These two results reiterate the strong trade-off we highlighted in Figure 3. The welfare of newborns in the history-independent allocation is 6.6% higher than the status quo economy. Although large, this is substantially less than the welfare gains from the constrained=efficient allocation.

[16]Appendix Figure F.3 shows that allowing consumption to also depend on previous location (i.e. one-period history dependence; Appendix C) yields an allocation that is nearly identical to the purely history-independent case. This suggests that to generate meaningful dynamic incentives, policy must depend on a longer span of migration history.

Table 3: Sensitivity to Calibrated Parameters

	b^C		b^{pop}		ΔGDP	$\Delta \text{Var}_L(\bar{C}_j)$
	DE	SP	DE	SP		
0. Baseline	0.81	0.39	0.39	6.06	+8.9%	-31.3%
1. No agglomeration	0.81	0.30	0.39	5.06	+8.0%	-59.8%
2. Congestion	0.81	0.23	0.39	3.87	+4.2%	-70.8%
3. Hand-to-mouth	0.82	0.41	0.38	5.20	+7.5%	-52.0%
4. Lower MPC	0.69	0.37	0.35	6.45	+9.3%	-33.0%
5. Lower migration elas.	0.81	0.40	0.39	5.47	+7.9%	-42.4%
6. Higher migration elas.	0.80	0.35	0.40	6.34	+10.2%	-25.5%
7. Higher risk aversion	0.81	0.31	0.38	5.39	+7.8%	-57.8%

Notes: This table compares the constrained-efficient allocation (SP) with the status quo economy (DE) for various alternative calibrations. The coefficient b^C is obtained from a linear regression of the form $\bar{C}_j = b^C\,(w/P)_j^{DE} + a^C + \epsilon_j$, where $\bar{C}_j$ is the state-level per capita consumption and $(w/P)_j^{DE}$ is the real wage in the status quo economy. Likewise, the coefficient b^{pop} is obtained from a linear regression of the form $pop_j = b^{pop}\,(w/P)_j^{DE} + a^{pop} + \epsilon_j$, where pop_j is state population. ΔGDP_j and $\Delta\text{Var}_L(\bar{C}_j)$ are changes in real GDP and population-weighted variance of state-level per capita consumption from the status quo economy to the constrained-efficient allocation (see the notes under Table 2 for precise definitions). Row 0 is our baseline economy. Row 1 considers an economy without agglomeration forces ($\alpha = 0$). Row 2 considers an economy with congestion externalities ($\alpha = -0.02$). Row 3 considers an economy with hand-to-mouth households ($\underline{a} \geq 0$). Row 4 considers a larger negative value for $\underline{a}$, which implies a low MPC calibration with a 5-year MPC of 0.5. Row 5 considers a lower migration elasticity ($\theta = 2$). Row 6 considers a higher migration elasticity ($\theta = 3$). Row 7 considers a higher risk aversion parameter ($\gamma = 1.2$). In all cases, we re-calibrate the parameters in Panel B of Table 1 to target the same moments.

6.6 Sensitivity to Calibrated Parameters

Table 3 presents results under alternative calibrations to our baseline economy. Rows 1 and 2 consider economies with either no net agglomeration forces ($\alpha = 0$) or negative agglomeration externalities (i.e. positive congestion effects; $\alpha = -0.02$). The qualitative patterns remain unchanged, although we find that the constrained-efficient allocation exhibits smaller GDP gains (ΔGDP) and a flatter profile of consumption (a larger negative value of $\Delta\text{Var}_L(\bar{C}_j)$). This is consistent with the interpretation that a lower α weakens the desirability of concentrating population in more productive regions.

Rows 3 and 4 explore alternative borrowing constraints in the status quo economy. Row 3 imposes a hand-to-mouth condition ($\underline{a} \geq 0$), while Row 4 allows for a larger negative asset limit, targeting a marginal propensity to consume (MPC) of 0.5, lower than the baseline MPC of 0.7. Even under this more conservative assumption about the borrowing constraint, the planner continues to smooth consumption more than in the status quo, reflecting the incomplete market of the status quo economy.

Rows 5 and 6 vary the migration elasticity, considering values lower ($\theta = 2$) and higher ($\theta = 3$) than the baseline. A higher θ results in a smaller change in the spatial variation of consumption ($\Delta \mathrm{Var}_L(\bar{C}_j)$). This is consistent with an interpretation that more elastic migration leads to less consumption smoothing, as evident from Proposition 2. Row 7 considers an economy with higher risk aversion. As expected, this leads to more equal consumption across space.

Overall, across all alternative calibrations, the planner's solution consistently features greater spatial equality in consumption and increased population concentration in more productive regions.

7 Transitions in Response to Aggregate Shocks

So far, we have focused on the steady-state. Now we introduce aggregate shocks to our baseline economy to study transition dynamics.

7.1 Introducing Aggregate Shocks

We analyze the transition dynamics in response to one-time shocks to technology. We assume that the aggregate shock arrives with probability $p > 0$. Let $x = 0, 1, \ldots$ denote the time elapsed since the arrival of the aggregate shock. If the shocks have not occurred yet, the technology evolves according to

$$f_{jt}(\{l_{kjt}\}, \{L_{kt}\}) = \begin{cases} f_j^0(\{l_{kjt}, L_{kt}\}) & \text{with prob. } p \\ f_j(\{l_{kjt}, L_{kt}\}) & \text{with prob. } 1 - p \end{cases}, \tag{56}$$

where f_j is the technology before the realization of the shock that is constant over time, and f_j^0 is the technology immediately after the realization of the shock. After the arrival of the aggregate shock, the technology is given by the deterministic sequence $\{f_j^x\}_{x=0}^{\infty}$, and we assume the sequence is convergent:

$$f_j^x \to f_j^{\infty} \quad \text{as} \quad x \to \infty. \tag{57}$$

Consequently, all aggregate variables (e.g. w_t^x) follow a convergent and deterministic sequence after the arrival of the aggregate shock. The arrival of the aggregate shock at $t + 1$ is announced at the end of time t before migration takes place.

The component planning problem for a household with promised utility v in location i in period t, before the realization of the aggregate shock, is

$$
\begin{aligned}
S_{it}(v) = & \max_{C_{it},\{v_{ijt+1},\mu^x_{ijt},v^0_{ijt+1},\mu^0_{ijt}\}} w_{it}\left(1+\alpha_{it}\right) - P_{it}C_{it} + (1-\omega)S_{it} \\
& + (1-p)\left[\frac{1}{R}\sum_j \mu_{ijt}S_{jt+1}(v_{ijt+1})\right] + p\left[\frac{1}{R}\sum_j \mu^0_{ijt}S^0_{jt+1}(v^0_{ijt+1})\right] \\
\text{s.t.} \quad & v = u_{it}(C_{it}) + \beta\omega\left\{(1-p)\left[\sum_j \mu_{ijt}v_{ijt+1} - \psi_{it}(\{\mu_{ijt}\})\right] + p\left[\sum_j \mu^0_{ijt}v^0_{ijt+1} - \psi_{it}(\{\mu^0_{ijt}\})\right]\right\} \\
& \{\mu_{ijt}\} \in \arg\max_{\{\tilde{\mu}_{ijt}\}}\left\{\sum_j \tilde{\mu}_{ijt}v_{ijt+1} - \psi_{it}(\{\tilde{\mu}_{ijt}\})\right\} \\
& \{\mu^0_{ijt}\} \in \arg\max_{\{\tilde{\mu}^0_{ijt}\}}\left\{\sum_j \tilde{\mu}^0_{ijt}v^0_{ijt+1} - \psi_{it}(\{\tilde{\mu}^0_{ijt}\})\right\},
\end{aligned}
\tag{58}
$$

where the variables without superscript 0 denote those before the arrival of the shock. The component planning problem for a newborn is likewise given by

$$
\begin{aligned}
S^n_{it} = & \max_{v^n_{it},\{v^n_{ijt+1},\mu^n_{ijt},v^{n,0}_{ijt+1},\mu^{n,0}_{ijt}\}} (1-p)\left[\Lambda_i v^n_{it} + \frac{1}{R}\sum_j \mu^n_{ijt}S_{jt+1}(v^n_{ijt})\right] \\
& + p\left[\Lambda_i v^{n,0}_{it} + \frac{1}{R}\sum_j \mu^{n,0}_{ijt}S^0_{jt+1}(v^{n,0}_{ijt})\right] \\
\text{s.t.} \quad & v^n_{it} = \beta\left[\sum_j \mu^n_{ijt}v^n_{ijt+1} - \psi_{it}(\{\mu^n_{ijt}\})\right], \quad v^{n,0}_{it} = \beta\left[\sum_j \mu^{n,0}_{ijt}v^{n,0}_{ijt+1} - \psi_{it}(\{\mu^{n,0}_{ijt}\})\right] \\
& \mu^n_{ijt} \in \arg\max_{\{\tilde{\mu}^n_{ijt}\}}\left\{\sum_j \tilde{\mu}^n_{ijt}v^n_{ijt+1} - \psi_{it}(\{\tilde{\mu}^n_{ijt}\})\right\} \\
& \mu^{n,0}_{ijt} \in \arg\max_{\{\tilde{\mu}^{n,0}_{ijt}\}}\left\{\sum_j \tilde{\mu}^{n,0}_{ijt}v^{n,0}_{ijt+1} - \psi_{it}(\{\tilde{\mu}^{n,0}_{ijt}\})\right\}.
\end{aligned}
\tag{59}
$$

After the realization of aggregate shock, the component planning problems are analogous to (31) and (34) indexed with superscript $x = 0, 1, 2, \ldots$ The Bellman equations for the transition dynamics of the status quo economy are modified in a similar way (see Appendix F.5).

A key challenge in studying the transition dynamics is the high dimensionality of the state space: with aggregate shocks, the full distribution of population over location and asset (in the status quo) or location and promised utility (in the constrained-efficient allocation) becomes state variables. Evaluating the net social surplus functions for all possible states is practically infeasible.

To address this challenge, we assume the aggregate shock occurs with arbitrarily small probability:

$$p \to 0. \tag{60}$$

This assumption dramatically simplifies our analysis for two reasons. First, it implies that the economy is in a deterministic steady-state before the realization of the aggregate shock. Second, it allows us to compute a first-order approximation of the transition path solely using the sequence-space Jacobian (Auclert et al. 2021) with respect to aggregate variables such as local wages and population size – significantly lower-dimensional objects. We describe the details of the computational algorithm in Appendices F.4 and F.5.

Importantly, this assumption is distinct from an "MIT shock", a one-time *unanticipated* shock (Mukoyama 2021). Here, the shock is anticipated, and the planner writes contingent plans in response to the shock. This distinction is important, as an "MIT shock" is ill-suited to study the dynamics of optimal policy responses as in our context. If shocks were unanticipated, the planner would need to re-optimize in response to the shock. However, allowing re-optimization introduces a time inconsistency problem: once migration decisions are realized, the planner has no incentive to fulfill any ex-ante plan for consumption front-loading or back-loading. This re-optimization therefore creates artificial dynamics unrelated to the aggregate shock itself.

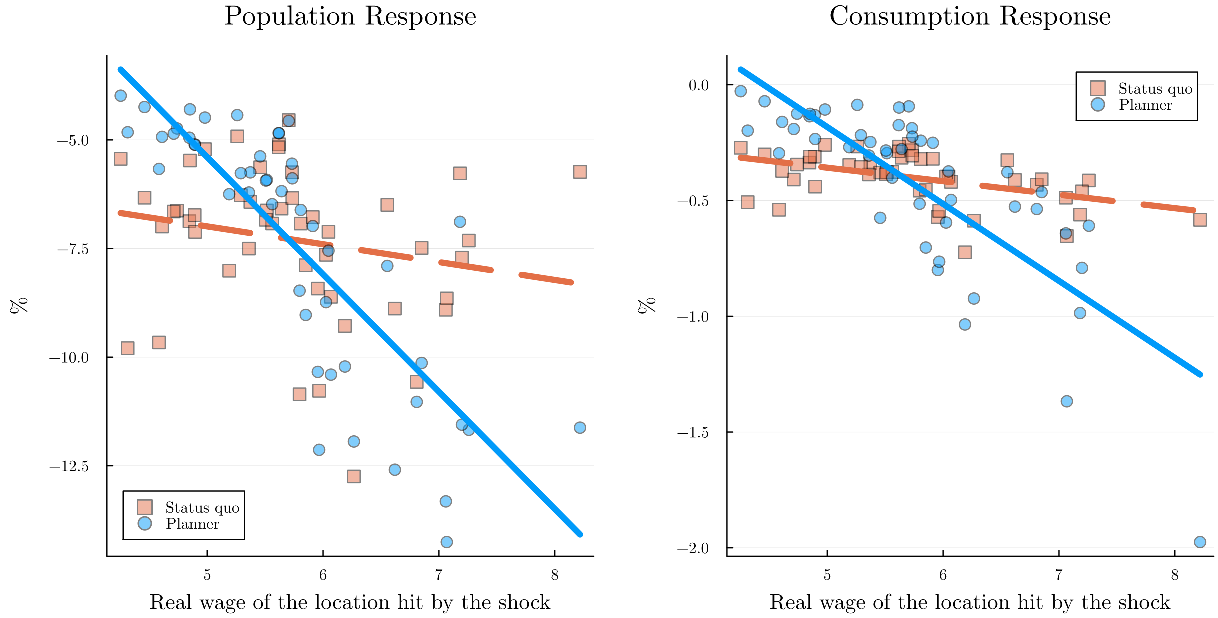
7.2 Long-Run Response to the Localized Productivity Shock

We consider a 1% negative permanent productivity shock to each of the 48 states, so there are 48 experiments in total. We first discuss how the steady-state (long-run) allocation changes in response to the shock. We discuss the transition dynamics from the old to the new steady-state in the next subsection.

The left panel of Figure 4 shows the long-run change in population in a shocked location against that location's real wage in the status quo economy. The red squares are the responses for the status quo economy, and the blue circles are those for the planner's solution. The linear fits for each economy are shown in the same color. In the status quo, the elasticity of population with respect to the shock is relatively flat across locations, averaging around 7.5%. In contrast, under the planner's solution (constrained-efficient allocation), the elasticity increases sharply with real wage: it is below 5% in the least productive locations but rises to nearly 15% in the most productive ones.

The right panel of Figure 4 shows the corresponding per capita consumption response. The pattern closely mirrors the population response. Under the constrained-efficient allocation, the elasticity of consumption with respect to the shock increases strongly with real wage, whereas it remains relatively constant in the status quo.

Figure 4: Long-run Impact of Localized Negative Productivity Shock



Notes: The left panel plots the population response (in percentage terms) of the location hit by the 1% negative productivity shock on the y-axis and the real wage in the same location in the status quo economy on the x-axis. The red squares are from the status quo economy, with the red dashed line as the best linear fit. The blue circles are from the planner's allocation, with the blue solid line as the best linear fit. The right panel plots the long-run pass-through from real wage to average consumption in the location hit by the shock, $\frac{d\ln C_i}{d\ln(w_i/P_i)}$, and is otherwise analogous to the left panel.

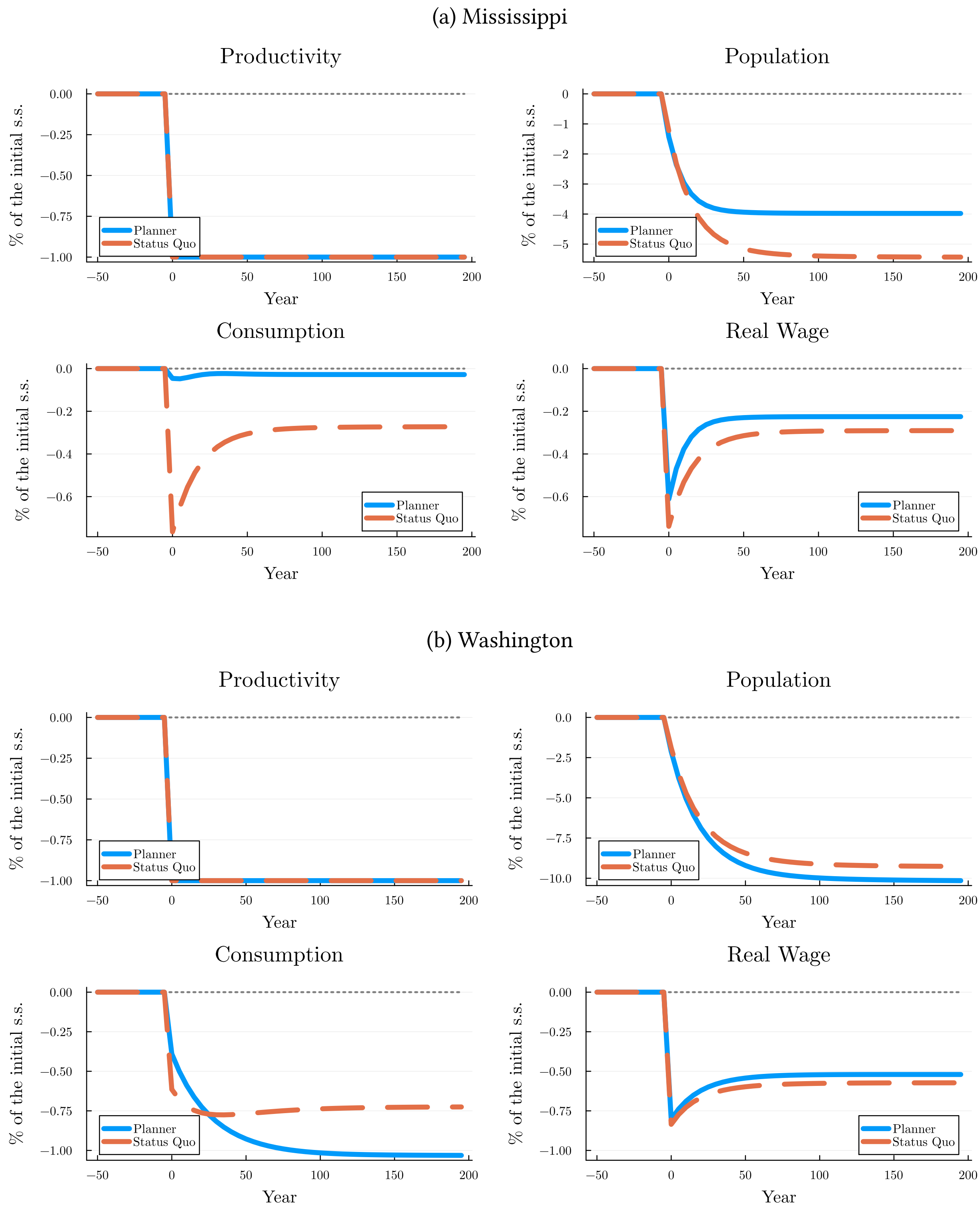
Why are the elasticities of local population and consumption with respect to local productivity significantly larger in more productive locations under the planner's solution? The key reason is that productive locations typically run a fiscal surplus, i.e. $w_j - P_jC_j > 0$. In such locations, consumption has more room to adjust, in proportional terms, than the real wage. In contrast, in locations with a fiscal deficit ($w_j - P_jC_j < 0$), consumption has less room to adjust, in proportional terms, than the real wage. As a result, shocks to surplus-generating locations generally lead to larger changes in consumption, and in turn, greater adjustments in population.[17]

7.3 Transition Dynamics

We now turn to the transition dynamics from the old to the new steady-state following the realization of a localized negative productivity shock. As examples, Figure 5a and 5b illustrate the transition path of Mississippi and Washington when each of them experiences a permanent 1% decline in productivity, respectively. In each plot, the blue solid line represents the planner's solution (the constrained-efficient allocation), and the red dashed line represents the status quo economy. "Productivity" panels show the productivity paths, which are the shock process we

[17]See Appendix E for more formal analytical comparative statics in a stylized setting.

Figure 5: Impulse Responses to a 1% Permanent Negative Productivity Shock



Notes: The figure shows the impulse responses of Mississippi and Washington to a 1% permanent negative productivity shock in each location. Panel (a) shows the response of Mississippi when Mississippi experiences the shock. Panel (b) shows the response of Washington when Washington experiences the shock. The blue solid line indicates the planner's solution (constrained-efficient allocation), and the red dashed line indicates the status quo. All responses are expressed as a percentage deviation from the initial steady-state. "Real wage" refers to w_{it}/P_{it} in both planner's solution and in the status quo.

feed in.

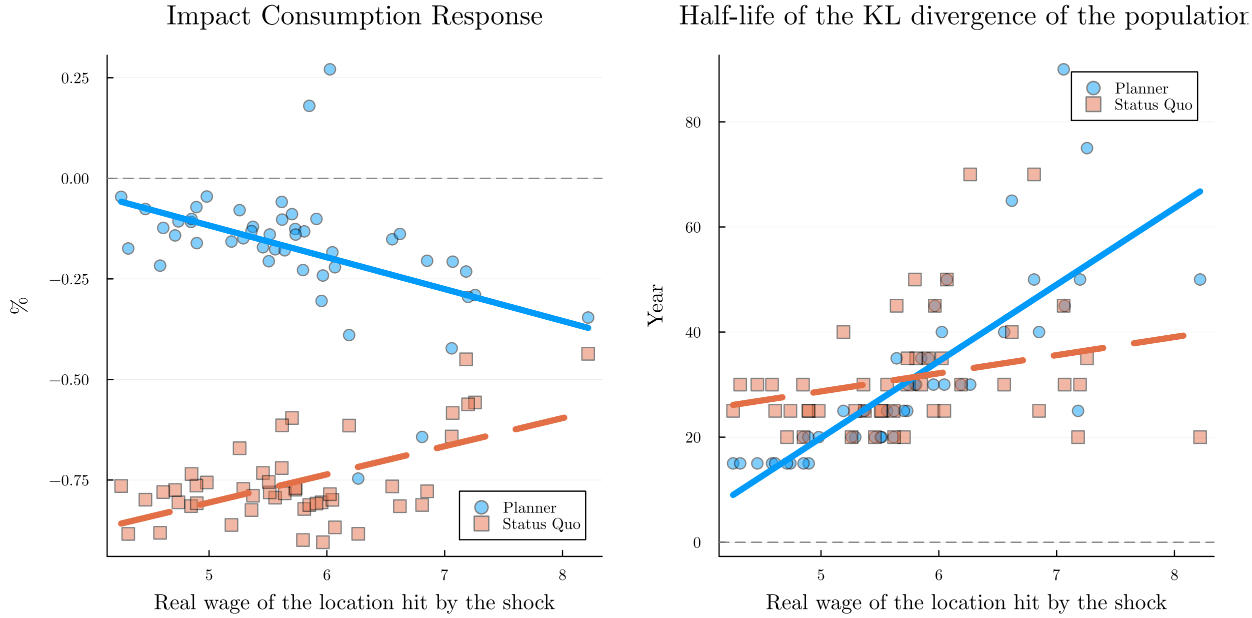
The bottom left panel in Figure 5a shows the consumption response in Mississippi. As discussed earlier, in the long run, the constrained-efficient allocation involves a smaller population decline in response to a shock in an unproductive state like Mississippi. In the status quo, the consumption response closely tracks the response of the real wage, shown in the bottom right panel. This is because most households in Mississippi are borrowing-constrained. Consumption drops abruptly and gradually recovers as households migrate out of Mississippi. In stark contrast, the constrained-efficient allocation achieves substantially better consumption smoothing, with an initial drop in consumption roughly equal to the long-term response. In other words, the constrained-efficient allocation features minimal transition dynamics in consumption relative to the status quo. This then leads to faster transition dynamics for population in the constrained-efficient allocation.

Figure 5b shows the response of Washington, a productive state, which contrasts with the response of Mississippi. First, in the status quo, households in Washington achieve better consumption smoothing than those in Mississippi, as they are less likely to be borrowing-constrained. After the initial drop, consumption immediately stabilizes near the new steady-state value. Second, consumption in the constrained-efficient allocation is substantially more front-loaded than the status quo. The consumption drops by a small amount initially and continues to decline over the next 100 years. This pattern reflects the planner's dual objective: to insure households against negative productivity shocks while encouraging long-run outmigration from affected locations. To support insurance, the planner minimizes the drop in initial consumption to smooth utility in the short run. But to incentivize outmigration over time, long-run consumption in the shocked location must fall. Third, this front-loaded and persistent consumption path, in turn, leads to a slower adjustment in population, as shown in the top right panel.

Figure 6 generalizes the patterns we highlight using Mississippi and Washington. The left panel plots the on-impact consumption response against the real wage in the shocked state. We see that the on-impact consumption response is substantially smaller in the planner's solution relative to the status quo, despite Figure 4 showing that the long-run consumption responses are similar on average. Interestingly, comparing with Figure 4, the difference between the initial and long-term consumption responses is greater in more productive locations. Therefore, the planner front-loads consumption profile more in productive locations.

The right panel depicts the speed of convergence of the population distribution, measured by the half-life of the Kullback–Leibler divergence. It shows that convergence is faster under the constrained-efficient allocation compared when the status quo when unproductive locations are hit by the shock. In contrast, convergence is slower in the constrained-efficient allocation when the productive locations are hit by the shock. This is consistent with a front-loaded dynamic

Figure 6: Response to 1% Permanent Negative Productivity Shock in Each State



Notes: The left panel plots the initial impact response on consumption in the shocked state to a 1% permanent negative productivity shock against the real wage of the shocked location in the status quo economy. The right panel plots the half-life of the Kullback–Leibler (KL) divergence of the population distribution in response to a 1% permanent negative productivity shock to each state against the real wage of the shocked location in the status quo economy. In both panels, the blue circles are the planner's solution, the red squares are the status quo economy, and the best linear fits for each case are shown in the same color.

consumption response in productive locations, which slows down the transitions, as well as little transition dynamics in consumption in unproductive locations, as discussed above.

Together, these patterns underscore the importance of dynamic incentives in optimal policy responses to aggregate shocks. Importantly, the planner's optimal response differs substantially between the short and long run, with systematic variation across locations depending on their pre-shock fundamentals.

8 Concluding Remarks

Many important real-life decisions are dynamic and discrete. This is why dynamic discrete choice models have been an extremely popular framework for studying various questions in labor, macro, industrial organization, international trade, and spatial economics. Despite its popularity, little is known about optimal policy in these environments. Methodologically, we provide a general framework for studying optimal policies in general equilibrium dynamic discrete models. Importantly, our framework does not impose any ad-hoc restrictions on the policy instruments. Substantively, we apply our framework to study optimal dynamic spatial policy in the US states. We find that dynamic incentives (i) alleviate the trade-off between spatial inequality and efficient population allocation in the long-run and (ii) critically shape the consumption and population response to localized technology shocks.

It goes without saying that our framework and results are not the final word on optimal dynamic spatial policy. In particular, we have assumed that the government can fully commit to future policies, which critically shapes the back-loading and front-loading of consumption profiles both in the steady-state and transition. We also have assumed a closed economy, but many important migration policies involve international migration flows. What if the government lacks commitment? What if the government faces international immigration flows? We plan to tackle these questions in future research.

References

Adão, R., A. Costinot, and D. Donaldson (2017): "Nonparametric Counterfactual Predictions in Neoclassical Models of International Trade," *American Economic Review*, 107, 633–689.

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 109, 659–684.

Ales, L. and C. Sleet (2022): "Optimal Taxation of Income-Generating Choice," *Econometrica*, 90, 2397–2436.

Allen, T. and D. Donaldson (2020): "Persistence and Path Dependence in the Spatial Economy," Working Paper 28059, National Bureau of Economic Research.

Anderson, S., A. De Palma, and J.-F. Thisse (1988): "A Representative Consumer Theory of the Logit Model," *International Economic Review*, 461–466.

Artuç, E., S. Chaudhuri, and J. McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100, 1008–1045.

Atkeson, A. and R. Lucas (1992): "On Efficient Distribution with Private Information," *The Review of Economic Studies*, 59, 427–453.

Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2021): "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models," *Econometrica*, 89, 2375–2408.

Auclert, A., M. Rognlie, L. Straub, and T. Tapak (2024): "When do Endogenous Portfolios Matter for HANK," *Working Paper*.

Baily, M. (1978): "Some Aspects of Optimal Unemployment Insurance," *Journal of Public Economics*, 10, 379–402.

Balke, N. and T. Lamadon (2022): "Productivity Shocks, Long-Term Contracts, and Earnings Dynamics," *American Economic Review*, 112, 2139–77.

Bewley, T. (1986): "Stationary monetary equilibrium with a continuum of independently fluctuating consumers," *Contributions to mathematical economics in honor of Gérard Debreu*, 79.

Bilal, A. and E. Rossi-Hansberg (2023): "Anticipating Climate Change Across the United States," Working Paper 31323, National Bureau of Economic Research.

Bloedel, A., R. V. Krishna, and O. Leukhina (2025): "Insurance and inequality with persistent private information," *Econometrica*, 93, 821–857.

BOEHM, J., E. FIZE, AND X. JARAVEL (2025): "Five facts about MPCs: Evidence from a randomized experiment," *American Economic Review*, 115, 1–42.

BURDETT, K. AND M. COLES (2003): "Equilibrium Wage-Tenure Contracts," *Econometrica*, 71, 1377–1404.

CALIENDO, L., M. DVORKIN, AND F. PARRO (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87, 741–835.

CALIENDO, L. AND F. PARRO (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 82, 1–44.

CARROLL, C. (2006): "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, 91, 312–320.

CHETTY, R. (2006): "A General Formula for the Optimal Level of Social Insurance," *Journal of Public Economics*, 90, 1879–1901.

COLARIETI, R., P. MEI, AND S. STANTCHEVA (2024): "The how and why of household reactions to income shocks," Tech. rep., National Bureau of Economic Research.

COSTINOT, A. AND A. RODRÍGUEZ-CLARE (2014): "Trade theory with numbers: Quantifying the consequences of globalization," in *Handbook of international economics*, Elsevier, vol. 4, 197–261.

DESMET, K. AND F. PARRO (2025): "Spatial Dynamics," Tech. rep., National Bureau of Economic Research.

DIAMOND, P. AND J. MIRRLEES (1978): "A model of social insurance with variable retirement," *Journal of Public Economics*, 10, 295–336.

DONALD, E., M. FUKUI, AND Y. MIYAUCHI (2025): "Unpacking Aggregate Welfare in a Spatial Economy," Working Paper 34075, National Bureau of Economic Research.

DVORKIN, M. (2023): "Heterogeneous Agents Dynamic Spatial General Equilibrium," *Working Paper*.

D'AMICO, L. AND M. ALEKSEEV (2024): "Capital Market Integration and Growth Across the United States," *Working Paper*.

EATON, J. AND S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

Eden, M. (2023): "The Cross-Sectional Implications of the Social Discount Rate," *Econometrica*, 91, 2065–2088.

Fagereng, A., M. Holm, and G. Natvik (2021): "MPC Heterogeneity and Household Balance Sheets," *American Economic Journal: Macroeconomics*, 13, 1–54.

Fajgelbaum, P. and C. Gaubert (2025): "Optimal Spatial Policies," Tech. rep., National Bureau of Economic Research.

Farhi, E. and I. Werning (2007): "Inequality and Social Discounting," *Journal of Political Economy*, 115, 365–402.

——— (2012): "Capital Taxation: Quantitative Explorations of the Inverse Euler Equation," *Journal of Political Economy*, 120, 398–445.

Fukui, M. (2020): "A theory of wage rigidity and unemployment fluctuations with on-the-job search," *Working Paper*.

Gaubert, C., P. Kline, and D. Yagan (2021): "Place-Based Redistribution," Working Paper 28337, National Bureau of Economic Research.

Giannone, E., Q. Li, N. Paixão, and X. Pang (2023): "Unpacking Moving," *Working Paper*.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003): "Optimal Indirect and Capital Taxation," *The Review of Economic Studies*, 70, 569–587.

Greaney, B. (2023): "Homeownership and the Distributional Effects of Uneven Regional Growth," *Working Paper*.

Greaney, B., A. Parkhomenko, and S. Van Nieuwerburgh (2025): "Dynamic Urban Economics," Tech. rep., National Bureau of Economic Research, Inc.

Guerreiro, J., S. Rebelo, and P. Teles (2023): "Foreign Residents and the Future of Global Cities," Tech. rep., National Bureau of Economic Research.

Hofbauer, J. and W. Sandholm (2002): "On the Global Convergence of Stochastic Fictitious Play," *Econometrica*, 70, 2265–2294.

Hopenhayn, H. and J. P. Nicolini (1997): "Optimal Unemployment Insurance," *Journal of Political Economy*, 105, 412–438.

Howard, G. and H. Shao (2022): "Internal Migration and the Microfoundations of Gravity," *Working Paper*.

Huggett, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control*, 17, 953–969.

Imrohoroğlu, A. (1989): "Cost of business cycles with indivisibilities and liquidity constraints," *Journal of Political economy*, 97, 1364–1383.

Kennan, J. and J. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79, 211–251.

Kleinman, B., E. Liu, and S. Redding (2023): "Dynamic Spatial General Equilibrium," *Econometrica*, 91, 385–424.

Kocherlakota, N. (2005): "Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation," *Econometrica*, 73, 1587–1621.

Kurnaz, M., M. Michelini, H. Özdenören, and C. Sleet (2023): "Tax Design in Dynamic Discrete Choice Economies," *Working Paper*.

Lhuillier, H. (2023): "Should I Stay or Should I Grow?" *Working Paper*.

Melo, P., D. Graham, and R. Noland (2009): "A Meta-Analysis of Estimates of Urban Agglomeration Economies," *Regional Science and Urban Economics*, 39, 332–342.

Mongey, S. and M. Waugh (2024): "Discrete choice, complete markets, and equilibrium," Tech. rep., National Bureau of Economic Research.

Mukoyama, T. (2021): "MIT shocks imply market incompleteness," *Economics Letters*, 198, 109666.

Orchard, J., V. Ramey, and J. Wieland (2023): "Micro MPCs and macro counterfactuals: the case of the 2008 rebates," Tech. rep., National Bureau of Economic Research.

O'Connor, D. (2024): "Revitalize or Relocate: Optimal Place Based Transfers for Local Recessions," *Working Paper*.

Peters, M. (2022): "Market Size and Spatial Growth—Evidence From Germany's Post-War Population Expulsions," *Econometrica*, 90, 2357–2396.

Prescott, E. and R. Townsend (1984): "General Competitive Analysis in an Economy with Private Information," *International Economic Review*, 1–20.

Rogerson, W. (1985): "Repeated Moral Hazard," *Econometrica*, 69–76.

Souchier, M. (2022): "The pass-through of productivity shocks to wages and the cyclical competition for workers," *Working Paper*.

Stantcheva, S. (2020): "Dynamic Taxation," *Annual Review of Economics*, 12, 801–831.

Veracierto, M. (2022): "Moral Hazard, Optimal Unemployment Insurance, and Aggregate Dynamics," *Working Paper*.

——— (2023): "Aggregate dynamics in Mirrlees economies: The case of persistent shocks," *Working Paper*.

Online Appendix for:

Optimal Dynamic Spatial Policy

Eric Donald   Masao Fukui   Yuhei Miyauchi

Contents

A Proofs and Mathematical Details

A.1 Lagrangian of the Recursive Planning Problem

Let P_{it}, w_{it}, and $w_{it}\alpha_{it}$ denote the Lagrange multipliers of (10), (11), and (12), respectively. Let $\mathcal{S}_t(\phi_t)$ denote the associated Lagrangian of the problem (7). It is given by

$$\begin{aligned}
\mathcal{S}_t(\phi_t) = & \max_{\{C_{it}(v),\ell^n_{it+1}(\epsilon),\ell_{it+1}(v,\epsilon),v^n_{it},v^n_{ijt+1}(\epsilon),v_{ijt+1}(v,\epsilon),l_{it},\phi_t,L_{it},P_{it},w_{it},\alpha_{it}\}} \sum_i \Lambda_i v^n_{it}\,(1-\omega)\int d\phi_{it}(v) \\
& + \sum_i P_{it}\left[f_{it}(\{l_{kit}\},\{L_{kt}\}) - \int C_{it}(v)d\phi_{it}(v)\right] \\
& + \sum_i w_{it}\left[\int d\phi_{it}(v) - \sum_j l_{ijt}\right] \\
& + \sum_i \alpha_{it}w_{it}\left[\int d\phi_{it}(v) - L_{it}\right] \\
& + \frac{1}{R}\mathcal{S}_{t+1}(\phi_{t+1})
\end{aligned} \tag{A.1}$$

subject to

$$v^n_{it} = \beta\mathbb{E}_{it}\Big[\sum_j \mathbb{I}[\ell^n_{it+1}(\epsilon_{it})=j]\left\{v^n_{ijt+1}(\epsilon_{it})+\epsilon_{ijt}\right\}\Big]] \tag{A.2}$$

$$v = u_{it}(C_{it}(v)) + \beta\omega\mathbb{E}_{it}\Big[\sum_j \mathbb{I}[\ell_{it+1}(v,\epsilon_{it})=j]\left\{v_{ijt}(v,\epsilon_{it})+\epsilon_{ijt}\right\}\Big] \tag{A.3}$$

and the law of motion of the distribution

$$\begin{aligned}
\phi_{jt+1}(v) = & \sum_i \omega\mathbb{E}_{it}\left[\phi_{it}(v^{-1}_{ijt+1}(v,\epsilon_{it}))\mathbb{I}[\ell_{it+1}(v^{-1}_{ijt+1}(v,\epsilon_{it}),\epsilon_{it})=j]\right] \\
& + (1-\omega)L_{it}\mathbb{E}_{it}\left[\mathbb{I}[v^n_{ijt+1}(\epsilon_{it})=v]\mathbb{I}[\ell^n_{it+1}(\epsilon_{it})=j]\right].
\end{aligned} \tag{A.4}$$

We now guess and verify that the value function takes the following form:

$$\mathcal{S}_t(\phi_t) = \sum_i \int S_{it}(v)d\phi_{it}(v) + D_t. \tag{A.5}$$

First, observe that the flow value in (A.1) is additively separable in ϕ_{it} and i. Second, under the guess, using (A.4), the continuation value can be rewritten as

$$\begin{aligned}\frac{1}{R}\mathcal{S}_{t+1}(\phi_{t+1}) = \frac{1}{R}\omega\sum_i\int\mathbb{E}_{it}\left[\sum_j\mathbb{I}[\ell_{it+1}(v,\epsilon)=j]S_{jt+1}(v_{ijt+1}(v))\right]d\phi_{it}(v)\\+\frac{1}{R}(1-\omega)\sum_i L_{it}\mathbb{E}_{it}\left[\sum_j\mathbb{I}[\ell^n_{it+1}(\epsilon)=j]S_{jt+1}(v^n_{ijt+1})\right]+\frac{1}{R}D_{t+1}.\end{aligned}\tag{A.6}$$

With these observations, it is immediate to see that the guess satisfies the Bellman equation (A.1) with $S_{it}(v)$ solving (18) and D_t solving (15).

A.2 Proof of Proposition 1

Let $\Xi_{it}(v)$ be the Lagrange multiplier on the promise-keeping constraint (19). The first-order condition with respect to $v_{ijt+1}(v,\epsilon_{it})$ is given by

$$\frac{1}{R}\partial_v S_{jt+1}(v_{ijt+1}(v,\epsilon_{it}))+\beta\Xi_{it}(v)=0.\tag{A.7}$$

From this expression, it is clear that $v_{ijt+1}(v,\epsilon_{it})$ does not depend on idiosyncratic preference shocks ϵ_{it}, i.e. $v_{ijt+1}(v,\epsilon_{it})=v_{ijt+1}(v)$. The first-order condition with respect to $C_{it}(v)$ is

$$P_{it}=u'_{it}(C_{it}(v))\Xi_{it}(v).\tag{A.8}$$

The envelope condition is

$$\partial_v S_{it}(v)=-\Xi_{it}(v).\tag{A.9}$$

Combining the above three expressions, we have

$$-\frac{1}{R}\frac{P_{jt+1}}{u'_{jt+1}(C_{jt+1}(v_{ijt+1}(v)))}+\beta\frac{P_{it}}{u'_{it}(C_{it}(v))}=0,\tag{A.10}$$

which is (22).

The location choice maximizes the Lagrangian:

$$\ell_{it}(v,\epsilon_{it})\in\arg\max_l\left\{\beta\omega\Xi_{it}[v_{ilt+1}(v)+\epsilon_{ilt}]+\frac{1}{R}S_{lt+1}(v_{ilt+1}(v))\right\}.\tag{A.11}$$

Substituting the expression for Ξ_{it} in (A.9) gives (23).

$\square$

A.3 Proof of Proposition 2

Let Ξ_{it} be the Lagrange multiplier on the promise-keeping constraint (32). The first-order conditions with respect to C_{it} and v_{imt+1} are

$$\Xi_{it} u'_{it}(C_{it}) = P_{it} \tag{A.12}$$

$$\frac{1}{R}\mu_{imt} S'_{mt+1}(v_{imt+1}) + \Xi_{it}\beta\mu_{imt} + \frac{1}{R}\sum_k \frac{\partial \mu_{ikt}}{\partial v_{imt+1}} S_{kt+1}(v_{ikt+1}) = 0 \tag{A.13}$$

The envelope condition is

$$S'_{it}(v) = -\Xi_{it}. \tag{A.14}$$

Combining the three equations,

$$-\frac{1}{R}\mu_{imt}\frac{P_{mt+1}}{u'_{mt+1}(C_{mt+1})} + \frac{P_{it}}{u'_{it}(C_{it})}\beta\mu_{imt} + \frac{1}{R}\sum_k \frac{\partial \mu_{ikt}}{\partial v_{imt+1}} S_{kt+1}(v_{kt+1}) = 0 \tag{A.15}$$

Rewriting the above equation,

$$\mu_{imt}\left[\beta R \frac{u'_{mt+1}(C_{mt+1})/P_{mt+1}}{u'_{it}(C_{it})/P_{it}} - 1\right] + \sum_k \frac{\partial \mu_{ikt}}{\partial C_{mt+1}} \frac{S_{kt+1}(v_{ikt+1})}{P_{mt+1}} = 0, \tag{A.16}$$

as desired.

□

A.4 Proof of Corollary 1

We divide both sides of the expression in Proposition 2 by $u_{jt+1}(C_{jt+1})/P_{jt+1}$ to obtain

$$\mu_{ijt}\left[\beta R \frac{1}{u'_{it}(C_{it})/P_{it}} - \frac{1}{u'_{jt+1}(C_{jt+1})/P_{jt+1}}\right] + \sum_k \frac{\partial \mu_{ikt}}{\partial C_{jt+1}} \frac{1}{u'_{jt+1}(C_{jt+1})} S_{kt+1}(v_{ikt+1}) = 0, \tag{A.17}$$

We further rewrite this as

$$\mu_{ijt}\left[\beta R \frac{1}{u'_{it}(C_{it})/P_{it}} - \frac{1}{u'_{jt+1}(C_{jt+1})/P_{jt+1}}\right] + \sum_k \frac{\partial \mu_{ikt}}{\partial u_{jt+1}} S_{kt+1}(v_{ikt+1}) = 0, \tag{A.18}$$

where $u_{jt+1} \equiv u_{jt+1}(C_{jt+1})$. With a slight change in notation, we can equivalently express a household's migration decisions as

$$V_{it} = \max_{\{\mu_{ikt}\}} \sum_{\mu_{ikt}} \mu_{ikt}[u_{kt}(C_{kt+1}) + \beta\omega V_{kt+1}] - \psi_{it}(\{\mu_{ikt}\}) \tag{A.19}$$

From this expression, uniformly increasing u_{jt+1} for all j would not affect the choice probability:

$$\sum_j \frac{\partial \mu_{ikt}}{\partial u_{jt+1}} = 0 \tag{A.20}$$

for all i, k.

Using this property and summing (A.18) across j, we have

$$\beta R \frac{1}{u'_{it}(C_{it})/P_{it}} - \sum_j \mu_{ijt} \frac{1}{u'_{jt+1}(C_{jt+1})/P_{jt+1}} + \sum_k S_{kt+1}(v_{ikt+1}) \underbrace{\sum_j \frac{\partial \mu_{ikt}}{\partial u_{jt+1}}}_{=0} = 0, \tag{A.21}$$

so that

$$\frac{1}{u'_{it}(C_{it})/P_{it}} = \frac{1}{\beta R} \mathbb{E}_{it}\left[\frac{1}{u'_{jt+1}(C_{jt+1})/P_{jt+1}}\right], \tag{A.22}$$

as desired.

□

A.5 Proof of Proposition 3

The planning problem in a recursive form is

$$\mathcal{S}_t(\{v_{it}, L_{it}\}) = \max_{\{v_{jt+1}, v^n_{it}, C_{it}, \mu_{ijt}\}} \sum_i [w_{it}(1+\alpha_{it}) - P_{it}C_{it}]L_{it} + (1-\omega)\sum_i \Lambda_i v^n_{it} L_{it} + \frac{1}{R}\mathcal{S}_{t+1}(\{v_{jt+1}, L_{jt+1}\}) \tag{A.23}$$

$$\text{s.t.} \qquad v_{it} = u_{it}(C_{it}) + \beta\omega\left[\sum_j \mu_{ijt}v_{jt+1} - \psi_{it}(\{\mu_{ijt}\})\right] \tag{A.24}$$

$$v^n_{it} = \beta\left[\sum_j \mu_{ijt}v_{jt+1} - \psi_{it}(\{\mu_{ijt}\})\right] \tag{A.25}$$

$$\mu_{ijt} \in \arg\max_{\{\mu_{ijt}\}}\left\{\sum_j \mu_{ijt}v_{jt+1} - \psi_{it}(\{\mu_{ijt}\})\right\} \tag{A.26}$$

$$L_{jt+1} = \sum_k L_{kt}\mu_{kjt} \tag{A.27}$$

Let $\varkappa_{it}L_{it}$ be the Lagrange multiplier on constraint (C.2). The first-order condition w.r.t. v_{jt+1} is

$$\frac{1}{R}\left[\frac{\partial \mathcal{S}_{t+1}}{\partial v_{jt+1}} + \sum_i\sum_k \frac{\partial \mu_{ikt}}{\partial v_{jt+1}} L_{it}\frac{\partial \mathcal{S}_{t+1}}{\partial L_{kt+1}}\right] + \beta\sum_i(\omega\varkappa_{it} + (1-\omega)\Lambda_i)L_{it}\mu_{ijt} = 0 \tag{A.28}$$

The first-order condition w.r.t. C_{jt} is

$$P_{jt} = \varkappa_{jt}u'_{jt}(C_{jt}) \tag{A.29}$$

The envelope conditions are

$$\frac{\partial \mathcal{S}_t}{\partial v_{jt}} = -\varkappa_{jt}L_{jt} \tag{A.30}$$

$$\frac{\partial \mathcal{S}_t}{\partial L_{jt}} = w_{jt}(1+\alpha_{jt}) - P_{jt}C_{jt} + (1-\omega)\Lambda_j v^n_{jt} + \frac{1}{R}\sum_k \mu_{jkt}\frac{\partial \mathcal{S}_{t+1}}{\partial L_{kt+1}} \tag{A.31}$$

Define $S_{jt} \equiv \frac{\partial \mathcal{S}_t}{\partial L_{jt}}$. Combining the above expressions,

$$\frac{1}{R}\left[-\frac{P_{jt+1}}{u'_{jt+1}(C_{jt+1})}L_{jt+1} + \sum_i\sum_k \frac{\partial \mu_{ikt}}{\partial v_{jt+1}}L_{it}S_{kt+1}\right] + \beta\sum_i\left(\omega\frac{P_{it}}{u'_{it}(C_{it})} + (1-\omega)\Lambda_i\right)L_{it}\mu_{ijt} = 0 \tag{A.32}$$

From the chain rule, $\frac{\partial \mu_{ikt}}{\partial v_{jt+1}} = \frac{\partial \mu_{ikt}}{\partial C_{jt+1}} \frac{1}{u'_{jt}(C_{jt})}$. Therefore, by multiplying both sides by $Ru'_{jt+1}(C_{jt+1})/P_{jt+1}$, the above equation is rewritten as

$$\sum_i L_{it} \left[\mu_{ijt} \left[\beta R \omega \frac{u'_{jt+1}(C_{jt+1})/P_{jt+1}}{u'_{it}(C_{it})/P_{it}} + \beta R(1-\omega)\Lambda_i \frac{u'_{jt+1}(C_{jt+1})}{P_{jt+1}} - 1 \right] + \sum_k \frac{\partial \mu_{ikt}}{\partial C_{jt+1}} \frac{S_{kt+1}}{P_{jt+1}} \right] = 0. \tag{A.33}$$

□

B Decentralization

We present one example of how the constrained-efficient allocation can be implemented when there are no private savings, either because households do not have access to the credit market or because the government bans private savings. The underlying environment remains the same as described in Section 2. Here we focus on explaining the market structure.

The households supply labor in each location i at wage w_{it}. The price of final goods in each location is P_{it}. Let $\ell_t \in \{1, \ldots, J\}$ denote the location of living at time t, and let ℓ^t denote the history of location of living of any household. The government sends transfers $T_t(\ell^t)$ as a function of history of living locations.

The household problem in a recursive form is

$$v_{jt}(\ell^t) = \max_{C_{jt}, \{\mu_{jkt}\}} u_{jt}(C_{jt}) + \beta \omega \left[\sum_k \mu_{jkt} v_{kt+1}(\{\ell^t, k\}) - \psi_{jt}(\{\mu_{jkt}\}) \right] \tag{B.1}$$

$$\text{s.t.} \quad P_{jt} C_{jt} = w_{jt} + T_t(\ell^t). \tag{B.2}$$

Let $C_{jt}(\ell^t)$ and $\mu_{jkt}(\ell^t)$ denote the policy functions associated with the above problem.

The firm takes prices and the population distribution $\{L_{kt}\}$ as given. The profit maximization problem of firm in location i is

$$\max_{\{l_{kjt}\}} P_{jt} f_{jt}(\{l_{kjt}\}, \{L_{kt}\}) - \sum_k w_{kt} l_{kj}. \tag{B.3}$$

Here, agglomeration/congestion forces are externalities that are not internalized by private agents. The government budget constraint is

$$\sum_{\ell^t} T_t(\ell^t) \Phi_t(\ell^t) = 0, \tag{B.4}$$

where $\Phi_t(\ell^t)$ denote the measure of households with history ℓ^t. The goods market clearing condition is

$$\int C_{jt}(\ell^t) d\Phi_t(\ell^t) = f_{jt}(\{l_{kjt}\}, \{L_{kt}\}). \tag{B.5}$$

The factor market clearing condition is

$$\sum_j l_{kjt} = \int \mathbb{I}[\ell_t = k] d\Phi_t(\ell^t) = L_{kt}. \tag{B.6}$$

The distribution evolves according to the following law of motion:

$$\Phi_{t+1}(\{\ell^t, k\}) = \mu_{\ell_t kt}(\ell^t)\Phi_t(\ell^t). \tag{B.7}$$

In the above decentralized equilibrium, an appropriate choice of the transfer system $T_t(\ell^t)$ implements the constrained-efficient allocation characterized in Proposition 2. To see this, first note that the continuation value in the constrained-efficient allocation only depends on the location of living in the next period and the promised utility. Therefore, given the initial location where each household is born, the promised value is only a function of the history of living locations. Let $v(\ell^t)$ denote the promised value with a history of living location ℓ^t. Then, $C_{\ell_t t}(v(\ell^t))$ is the consumption of households currently in location ℓ_t with a history ℓ^t in the constrained-efficient allocation.

Consider the following transfer system:

$$T_t(\ell^t) = P_{\ell_t t} C_{\ell_t t}(v(\ell^t)) - w_{\ell_t t}. \tag{B.8}$$

From the budget constraint, it is immediate to see that such a transfer system implements the constrained-efficient allocation as long as $\{P_{jt}, w_{jt}\}$ in the decentralized equilibrium coincide with those in the constrained-efficient allocation. To see why, the migration probabilities are identical given $\{P_{jt}, w_{jt}\}_j$. The optimality conditions of (B.5) is identical to (16) in the constrained-efficient allocation. The market clearing conditions and the evolution of the distribution are identical in both economies by construction. Finally, the government budget (B.4) is satisfied by Walras' law. Given that all the conditions in the two economies coincide, the transfer scheme (B.8) implements the constrained-efficient allocation as a decentralized equilibrium.

C One-Period-History Constrained-Efficient Allocation

We consider the case where the planner specifies the consumption for each origin, destination, and time C_{ijt}. Note that this is an intermediate case between fully-history-dependent constrained-efficient allocation in Section 4.2 and the history-independent constrained-efficient allocation in Section 4.3.

The planning problem in a recursive form is

$$\mathcal{S}_t(\{v_{kit}, L_{kit}\}) = \max_{\{v_{it}^n, \{C_{kit}\}, \{v_{ijt+1}, \mu_{ijt}\}\}} \sum_i [w_{it}(1+\alpha_{it})L_{it} - P_{it}\sum_k C_{kit}L_{kit}] + (1-\omega)\sum_i \Lambda_i v_{it}^n L_{it} + \frac{1}{R}\mathcal{S}_{t+1}(\{v_{ijt+1}, L_{ijt+1}\}) \tag{C.1}$$

$$\text{s.t.} \quad v_{kit} = u_{it}(C_{kit}) + \beta\omega\left[\sum_j \mu_{ijt}v_{ijt+1} - \psi_{it}(\{\mu_{ijt}\})\right] \tag{C.2}$$

$$v_{it}^n = \beta\left[\sum_j \mu_{ijt}v_{ijt+1} - \psi_{it}(\{\mu_{ijt}\})\right] \tag{C.3}$$

$$\mu_{ijt} \in \arg\max_{\{\mu_{ijt}\}}\left\{\sum_j \mu_{ijt}v_{ijt+1} - \psi_{it}(\{\mu_{ijt}\})\right\} \tag{C.4}$$

$$L_{ijt+1} = L_{it}\mu_{ijt} \tag{C.5}$$

$$L_{it} = \sum_k L_{kit} \tag{C.6}$$

Let $\varkappa_{kit}L_{kit}$ be the Lagrange multiplier on constraint (C.2). The first-order condition w.r.t. v_{ijt+1} is

$$\frac{1}{R}\left[\frac{\partial\mathcal{S}_{t+1}}{\partial v_{ijt+1}} + \sum_k \frac{\partial\mu_{ikt}}{\partial v_{ijt+1}}L_{it}\frac{\partial\mathcal{S}_{t+1}}{\partial L_{ikt+1}}\right] + \beta\sum_k(\omega\varkappa_{kit} + (1-\omega)\Lambda_i)L_{kit}\mu_{ijt} = 0 \tag{C.7}$$

The first-order condition w.r.t. C_{ijt} is

$$P_{jt} = \varkappa_{ijt}u_{jt}'(C_{ijt}) \tag{C.8}$$

The envelope conditions are

$$\frac{\partial\mathcal{S}_t}{\partial v_{ijt}} = -\varkappa_{ijt}L_{ijt} \tag{C.9}$$

$$\frac{\partial\mathcal{S}_t}{\partial L_{ijt}} = w_{jt}(1+\alpha_{jt}) - P_{jt}C_{ijt} + (1-\omega)\Lambda_j v_{jt}^n + \frac{1}{R}\sum_k \mu_{jkt}\frac{\partial\mathcal{S}_{t+1}}{\partial L_{jkt+1}} \tag{C.10}$$

Define $S_{ijt} \equiv \frac{\partial \mathcal{S}_t}{\partial L_{ijt}}$. Combining the above expressions,

$$\frac{1}{R}\left[-\frac{P_{jt+1}}{u'_{jt+1}(C_{ijt+1})}L_{ijt+1}+\sum_k \frac{\partial \mu_{ikt}}{\partial v_{ijt+1}}L_{it}S_{ikt+1}\right]+\beta\sum_k\left(\omega\frac{P_{it}}{u'_{it}(C_{kit})}+(1-\omega)\Lambda_i\right)L_{kit}\mu_{ijt}=0 \tag{C.11}$$

From the chain rule, $\frac{\partial \mu_{ikt}}{\partial v_{ijt+1}} = \frac{\partial \mu_{ikt}}{\partial C_{ijt+1}}\frac{1}{u'_{jt+1}(C_{ijt+1})}$. Therefore, by multiplying both hand side by $Ru'_{jt+1}(C_{ijt+1})\frac{1}{P_{jt+1}}\frac{1}{L_{it}}$, above equation is rewritten as

$$\mu_{ijt}\left[\beta R\omega\frac{u'_{jt+1}(C_{ijt+1})}{P_{jt+1}}\sum_k\frac{L_{kit}}{L_{it}}\frac{P_{it}}{u'_{it}(C_{kit})}+\beta R(1-\omega)\Lambda_i\frac{u'_{jt+1}(C_{ijt+1})}{P_{jt+1}}-1\right]+\sum_k\frac{\partial\mu_{ikt}}{\partial C_{ijt+1}}\frac{S_{ikt+1}}{P_{jt+1}}=0. \tag{C.12}$$

D Extensions

D.1 Capital Accumulation

We now introduce capital accumulation in the baseline model. Assume that in each location j, there is a capital stock denoted as K_{jt}. The production function of the final goods at location j is now given by

$$Y_{jt} = f_{jt}(\{l_{kjt}, k_{kjt}, L_{kt}\}), \tag{D.1}$$

where k_{kjt} denotes the use of capital stock from location k in location j. The capital stock in location j at time t depreciates at rate δ_{jt}, but the final goods in location j can be invested into the capital stock in the same location. The law of motion of capital stock in location j is

$$K_{jt+1} = K_{jt}(1-\delta_{jt}) + I_{jt}, \tag{D.2}$$

where I_{jt} is the investment. The investment incurs the adjustment cost of the form $\Psi_{jt}(I_{jt}, K_{jt})$ in the units of final goods in location j. The capital market clearing condition is

$$\sum_k k_{jkt} = K_{jt}. \tag{D.3}$$

The goods market clearing condition in location j is modified as

$$\int_0^1 C_{jt}(h)dh + I_{jt} + \Psi_{jt}(I_{jt}, K_{jt}) = f_{jt}\left(\{l_{kjt}, k_{kjt}, L_{kt}\}\right). \tag{D.4}$$

The rest of the environment remains unchanged.

In this environment, there is no change in the component planning problem (31) and (34). The only change comes from D_t in (14). Let r_{jt} be the Lagrange multiplier on (D.3). The term D_t now includes the distribution of capital in each location as a state variable and is given by

$$D_t(\{K_{jt}\}) = \max_{\{l_{ijt},k_{ijt},L_{it},K_{jt+1}\}} \sum_i P_{it} f_{it}\left(\{l_{kit}, k_{kit}, L_{kt}\}\right) - \sum_i w_{it} \sum_j l_{ijt} \tag{D.5}$$

$$- \sum_i \alpha_{it} w_{it} L_{it} - \sum_i r_{it} K_{it} - \sum_i r_{it} \sum_j k_{ijt} - \sum_i P_{it}\left[I_{it} + \Psi_{it}(I_{it}, K_{it})\right] \tag{D.6}$$

$$+ \frac{1}{R} D_{t+1}(\{K_{jt+1}\}). \tag{D.7}$$

$$\text{s.t.} \quad K_{jt+1} = K_{jt}(1-\delta_{jt}) + I_{jt}. \tag{D.8}$$

The first-order conditions with respect to l_{ijt} and L_{it} remain essentially the same as in the main text:

$$P_{it} \frac{\partial f_{it}\left(\{l_{kit}, k_{kit}, L_{kt}\}\right)}{\partial l_{kit}} = w_{kt} \tag{D.9}$$

$$\sum_i P_{it} \frac{\partial f_{it}\left(\{l_{kit}, k_{kit}, L_{kt}\}\right)}{\partial L_{kt}} = \alpha_{kt} w_{kt}. \tag{D.10}$$

The optimality condition for the spatial allocation of capital services k_{ijt} is given by

$$P_{it} \frac{\partial f_{it}\left(\{l_{kit}, k_{kit}, L_{kt}\}\right)}{\partial k_{kit}} = r_{kt}. \tag{D.11}$$

The first-order condition with respect to investment I_{jt} is

$$P_{jt}\left(1 + \partial_I \Psi_{jt}(I_{jt}, K_{jt})\right) = \frac{\partial D_{t+1}(\{K_{jt+1}\})}{\partial K_{jt+1}}. \tag{D.12}$$

The envelope condition is

$$\frac{\partial D_t(\{K_{jt}\})}{\partial K_{jt}} = r_{jt} - P_{jt} \partial_K \Psi_{jt}(I_{jt}, K_{jt}) + (1-\delta) \frac{1}{R} \frac{\partial D_{t+1}(\{K_{jt+1}\})}{\partial K_{jt+1}}. \tag{D.13}$$

Let

$$q_{jt} \equiv \frac{\partial D_t(\{K_{jt}\})}{\partial K_{jt}} \tag{D.14}$$

be the "marginal q" of the capital stock in location j at time t. Using this expression, we can equivalently write (D.12) and (D.13) as

$$P_{jt}\left(1+\partial_I \Psi_{jt}(I_{jt},K_{jt})\right)=q_{jt+1} \tag{D.15}$$

and

$$q_{jt}=r_{jt}-P_{jt}\partial_K \Psi_{jt}(I_{jt},K_{jt})+(1-\delta)\frac{1}{R}q_{jt+1}. \tag{D.16}$$

Therefore, optimal investment follows the prescription of the q-theory of investment.

D.2 Ex-Ante Heterogeneous Household Types

In the baseline model, we have assumed that households are ex-ante homogeneous. We now consider an extension of the baseline model to multiple ex-ante heterogeneous household types.

There are M heterogeneous household types (e.g. race, skills, or gender). Each household dynasty h belongs to one of the types indexed by $\theta \in \{\theta_1,\ldots,\theta_M\}$, each of which has a mass ℓ^{θ}. We allow arbitrary heterogeneity across households with respect to θ, including preferences, location preference shock distributions, and death probabilities. When a household of type θ dies, they are replaced by a newborn of the same type, so the mass of type θ remains fixed at ℓ^{θ}. Importantly, we assume the ex-ante types are observable to the planner.

The technology to produce the final goods consumed by the household θ in location j at time t is

$$Y_{jt}^{\theta}=f_{jt}^{\theta}(\{l_{kjt}^{\tilde{\theta},\theta},L_{kt}^{\tilde{\theta}}\}), \tag{D.17}$$

where $l_{kjt}^{\tilde{\theta},\theta}$ denotes the labor services of type $\tilde{\theta}$ shipped from location k to j used to produce final goods for consumption goods of type θ, and L_{it}^{θ} is the population size of households of type θ in location i at time t. Here, we allow for agglomeration/congestion forces to depend arbitrarily on the population size of different household types.

The planner's objective is to maximize the following social welfare function:

$$\mathcal{W}_0=\sum_{t=0}^{\infty}\frac{1}{R^t}\sum_{\theta=\theta_1}^{\theta_M}\sum_{i=1}^{J}\Lambda_i^{\theta}v_{it}^{n,\theta}(1-\omega^{\theta})L_{it}^{\theta}, \tag{D.18}$$

where Λ_i^{θ} is the welfare weight attached to household of type θ born in location i.

Let φ_j^{θ} be the distribution over promised utility for households of type θ living in location j.

The goods market clearing condition for consumption goods of type θ is

$$\int C^{\theta}_{jt}(v)d\phi^{\theta}_{j}(v) = f^{\theta}_{jt}(\{l^{\tilde{\theta},\theta}_{kjt}, L^{\tilde{\theta}}_{kt}\}). \tag{D.19}$$

The labor market clearing condition for type θ is

$$\sum_{k,\tilde{\theta}} l^{\theta,\tilde{\theta}}_{jkt} = \int \mathbb{I}[\ell^{\theta}_{t}(v,\epsilon) = j]d\phi^{\theta}_{j}(v)dG_{jt}(\epsilon), \tag{D.20}$$

and the following equation dictates the agglomeration forces:

$$\int \mathbb{I}[\ell^{\theta}_{t}(v,\epsilon) = j]d\phi^{\theta}_{j}(v)dG_{jt}(\epsilon) = L^{\theta}_{jt}. \tag{D.21}$$

The evolution of distribution is

$$\begin{aligned}\phi^{\theta}_{jt+1}(v) = \sum_{i}\omega^{\theta}\mathbb{E}^{\theta}_{it}\left[\phi^{\theta}_{it}({v^{\theta}_{ijt+1}}^{-1}(v,\epsilon^{\theta}_{it}))\mathbb{I}[\ell^{\theta}_{it+1}({v^{\theta}_{ijt+1}}^{-1}(v,\epsilon^{\theta}_{it}),\epsilon^{\theta}_{it}) = j]\right] \\ + (1-\omega^{\theta})L^{\theta}_{it}\mathbb{E}^{\theta}_{it}\left[\mathbb{I}[v^{n,\theta}_{ijt+1}(\epsilon^{\theta}_{it}) = v]\mathbb{I}[\ell^{n,\theta}_{it+1}(\epsilon^{\theta}_{it}) = j]\right].\end{aligned} \tag{D.22}$$

Given this specification of the environment, the value function D_t in (15) in the main text is now replaced by

$$D_t = \max_{\{l^{\tilde{\theta},\theta}_{ijt}, L^{\theta}_{i}\}} \sum_{i,\theta} P^{\theta}_{it} f^{\theta}_{it}(\{l^{\tilde{\theta},\theta}_{kit}, L^{\tilde{\theta}}_{kt}\}) - \sum_{i,\theta} w^{\theta}_{it} \sum_{j,\tilde{\theta}} l^{\theta,\tilde{\theta}}_{ijt} - \sum_{i,\theta} \alpha^{\theta}_{it} w^{\theta}_{it} L^{\theta}_{it} + \frac{1}{R} D_{t+1}, \tag{D.23}$$

where P^{θ}_{jt}, w^{θ}_{jt}, and α^{θ}_{jt} are Lagrange multipliers on (D.19), (D.20), and (D.21). The first-order optimality conditions are

$$P^{\theta}_{it}\frac{\partial f^{\theta}_{it}(\{l^{\tilde{\theta},\theta}_{kit}, L^{\tilde{\theta}}_{kt}\})}{\partial l^{\tilde{\theta},\theta}_{kit}} = w^{\theta}_{kt} \tag{D.24}$$

$$\sum_{\hat{\theta}}\sum_{i} P^{\hat{\theta}}_{it}\frac{\partial f^{\hat{\theta}}_{it}(\{l^{\tilde{\theta},\hat{\theta}}_{kit}, L^{\tilde{\theta}}_{kt}\})}{\partial L^{\theta}_{kt}} = \alpha^{\theta}_{kt} w^{\theta}_{kt}. \tag{D.25}$$

The component planning problems are essentially the same as (31) except that now everything is indexed by θ:

$$S^{\theta}_{it}(v) = \max_{C^{\theta}_{it},\{v^{\theta}_{ijt+1},\mu^{\theta}_{ijt}\}} w^{\theta}_{it}\left(1+\alpha^{\theta}_{it}\right) - P^{\theta}_{it}C^{\theta}_{it} + (1-\omega^{\theta})S^{n,\theta}_{it} + \frac{1}{R}\omega^{\theta}\sum_j \mu^{\theta}_{ijt}S^{\theta}_{jt+1}\left(v^{\theta}_{ijt+1}\right) \tag{D.26}$$

$$\text{s.t.} \qquad v = u^{\theta}_{it}(C^{\theta}_{it}) + \beta^{\theta}\omega^{\theta}\left[\sum_j \mu^{\theta}_{ijt}v^{\theta}_{ijt+1} - \psi^{\theta}_{it}(\{\mu^{\theta}_{ijt}\})\right] \tag{D.27}$$

$$\{\mu^{\theta}_{ijt}\} \in \arg\max_{\{\tilde{\mu}^{\theta}_{ijt}\}} \left\{\sum_j \tilde{\mu}^{\theta}_{ijt}v^{\theta}_{ijt+1} - \psi^{\theta}_{it}(\{\tilde{\mu}^{\theta}_{ijt}\})\right\} \tag{D.28}$$

and the following replaces (34):

$$S^{n,\theta}_{it} = \max_{v^{n,\theta}_{it},\{v^{n,\theta}_{ijt+1},\mu^{n,\theta}_{ijt}\}} \Lambda^{\theta}_i v^{n,\theta}_i + \frac{1}{R}\sum_j \mu^{n,\theta}_{ijt}S^{\theta}_{jt+1}(v^{n,\theta}_{ijt}) \tag{D.29}$$

$$\text{s.t.} \qquad v^{n,\theta}_{it} = \beta^{\theta}\sum_j \left[\mu^{n,\theta}_{ijt}v^{n,\theta}_{ijt} - \psi^{\theta}_{it}(\{\mu^{n,\theta}_{ijt}\})\right] \tag{D.30}$$

$$\{\mu^{n,\theta}_{ijt}\} \in \arg\max_{\{\tilde{\mu}^{n,\theta}_{ijt}\}} \left\{\sum_j \tilde{\mu}^{n,\theta}_{ijt}v^{n,\theta}_{ijt} - \psi^{\theta}_{it}(\{\tilde{\mu}^{n,\theta}_{ijt}\})\right\}. \tag{D.31}$$

The following formula is an analogue of Proposition 2:

$$\mu^{\theta}_{ijt}\left[\beta^{\theta}R\frac{u^{\theta\prime}_{jt+1}(C^{\theta}_{jt+1})/P^{\theta}_{jt+1}}{u^{\theta\prime}_{it}(C^{\theta}_{it})/P^{\theta}_{it}} - 1\right] + \underbrace{\sum_k \frac{\partial\mu^{\theta}_{ikt}}{\partial C^{\theta}_{jt+1}}\frac{S^{\theta}_{kt+1}(v^{\theta}_{ikt+1})}{P^{\theta}_{jt+1}}}_{\equiv\,\xi^{\theta}_{ijt}} = 0. \tag{D.32}$$

D.3 Lagged Agglomeration/Congestion Forces

In the baseline model, we assumed that agglomeration/congestion forces only depend on the contemporaneous population size distribution. We now allow for agglomeration/congestion forces to depend on lagged population size distributions.

The only modification is that now the production function takes the following form:

$$Y_{jt} = f_{jt}\left(\{l_{kjt}, L^t_{kt}, L^t_{kt-1}, \dots, L^t_{kt-T^L}\}\right), \tag{D.33}$$

which replaces (3) in the main text. Here L^t_{kt-s} denotes the population size in location k at time $t-s$ that enters as an agglomeration/congestion force for production at time t. The lagged population sizes are defined as

$$L^t_{jt-s} = \int d\phi_{jt-s}(v) \quad \text{for } s = 1, \ldots, T^L. \tag{D.34}$$

Let $\alpha^t_{jt-s} w_{jt}$ be the Lagrange multiplier on the above equation.

Given all these modifications, the value function D_t in (15) in the main text is now replaced by

$$D_t = \max_{\{\{l_{ijt}\}, L_{it}\}} \sum_i P_{it} f_{jt}\Big(\{l_{kjt}, L^t_{kt}, L^t_{kt-1}, \ldots, L^t_{kt-T^L}\}\Big) - \sum_i w_{it} \sum_j l_{ijt} \tag{D.35}$$

$$- \sum_{s=0}^{T^L} \sum_i \alpha^t_{it-s} w_{it} L^t_{it-s} + \frac{1}{R} D_{t+1}. \tag{D.36}$$

The first-order optimality conditions are

$$P_{it} \frac{\partial f_{it}}{\partial l_{kit}} = w_{kt} \tag{D.37}$$

$$\sum_i P_{it} \frac{\partial f_{it}}{\partial L_{kt-s}} = w_{kt-s} \alpha^t_{kt-s}. \tag{D.38}$$

The net social surplus function $S_{it}(v)$ now takes into account that increasing the current population size of a location changes production technologies in the future:

$$S_{it}(v) = \max_{C_{it}, \{v_{ijt+1}, \mu_{ijt}\}} w_{it} + \textstyle\sum_{s=0}^{T^L} \dfrac{1}{R^s} w_{it+s} \alpha^{t+s}_{it} - P_{it} C_{it} + (1-\omega) S^n_{it} + \dfrac{1}{R} \omega \displaystyle\sum_j \mu_{ijt} S_{jt+1}\big(v_{ijt+1}\big)$$

$$\text{s.t.} \qquad v = u_{it}(C_{it}) + \beta\omega \left[\sum_j \mu_{ijt} v_{ijt+1} - \psi_{it}(\{\mu_{ijt}\}) \right]$$

$$\{\mu_{ijt}\} \in \arg\max_{\{\tilde{\mu}_{ijt}\}} \left\{ \sum_j \tilde{\mu}_{ijt} v_{ijt+1} - \psi_{it}(\{\tilde{\mu}_{ijt}\}) \right\}.$$

The value function for newborns $S^n_{it}(v)$ remains unchanged and is given by (34).

E Optimal Response to a Productivity Shock in a Stylized Environment

Consider a special case of our baseline model with the following features: (i) the economy is static, $\beta \to 0$; (ii) there are two locations, $i, j \in \{1, 2\}$; (iii) there are no agglomeration externalities,

$\alpha_{jt} = 0$ for all j, t; (iv) preference shocks are given by a type-I extreme value distribution (55); and (v) the social welfare function has equal weights across locations, $\Lambda_j = 1$ for all j. At the beginning of the period, households make location choice decisions. As demonstrated by Donald, Fukui, and Miyauchi (2025), the optimal spatial transfer policy in this static spatial economy must satisfy

$$u_j'(C_j) - P_j = \sum_{i=1}^{2} \frac{\partial \mu_i}{\partial C_j} [P_j C_j - w_j] \tag{E.1}$$

for $j = 1, 2$. Under the logit discrete choice system, we have

$$\frac{\partial \mu_i}{\partial C_j} = \begin{cases} \theta \mu_j (1 - \mu_j) u_j'(C_j) & \text{for } i = j \\ -\theta \mu_j \mu_i u_j'(C_j) & \text{for } i \neq j \end{cases}. \tag{E.2}$$

Substituting this expression into the above expression, we have

$$1 - \frac{P_1}{u_1'(C_1)} = \theta \mu_1 \mu_2 \left([P_1 C_1 - w_1] - [P_2 C_2 - w_2] \right) \tag{E.3}$$

$$1 - \frac{P_2}{u_2'(C_2)} = \theta \mu_1 \mu_2 \left([P_2 C_2 - w_2] - [P_1 C_1 - w_1] \right). \tag{E.4}$$

We consider baseline technology and changes in technology such that there is a change in w_1 only, so price levels $\{P_j\}_{j=1}^2$ and wage in location 2 w_2 are left unchanged. In this case, perturbations of (E.3) and (E.4) gives

$$\begin{aligned} -\frac{P_1}{u_1'(C_1)} \gamma_1 d \ln C_1 &= \theta \mu_1 \mu_2 \left(P_1 C_1 d \ln C_1 - w_1 d \ln w_1 - P_2 C_2 d \ln C_2 \right) \\ &\quad + \theta (\mu_2 - \mu_1) \left([P_1 C_1 - w_1] - [P_2 C_2 - w_2] \right) d\mu_1 \end{aligned} \tag{E.5}$$

$$\begin{aligned} -\frac{P_2}{u_2'(C_2)} \gamma_2 d \ln C_2 &= \theta \mu_1 \mu_2 \left(P_2 C_2 d \ln C_2 + w_1 d \ln w_1 - P_1 C_1 d \ln C_1 \right) \\ &\quad + \theta (\mu_2 - \mu_1) \left([P_2 C_2 - w_2] - [P_1 C_1 - w_1] \right) d\mu_1, \end{aligned} \tag{E.6}$$

where $\gamma_i \equiv u_i''(C_i) C_i / u'(C_i)$ is the relative risk aversion of location i at the perturbation point. We further seek to simplify the above expressions by considering parameters such that $\mu_2 = \mu_1$.

In this case, we can solve for $d\ln C_1$ as

$$d\ln C_1 = \frac{\frac{\theta\mu_1\mu_2 w_1}{\theta\mu_1\mu_2 P_1 C_1 + \frac{P_1}{u_1'(C_1)}\gamma_1}\left\{\frac{\frac{P_2}{u_2'(C_2)}\gamma_2}{\theta\mu_1\mu_2 P_2 C_2 + \frac{P_2}{u_2'(C_2)}\gamma_2}\right\}}{1 + \frac{\theta\mu_1\mu_2 P_2 C_2}{\theta\mu_1\mu_2 P_1 C_1 + \frac{P_1}{u_1'(C_1)}\gamma_1}\frac{\theta\mu_1\mu_2 P_1 C_1}{\theta\mu_1\mu_2 P_2 C_2 + \frac{P_2}{u_2'(C_2)}\gamma_2}} d\ln w_1. \tag{E.7}$$

With log utility, $u_j(C_j) = \ln C_j$ for $j = 1, 2$, which we assume in the quantitative analysis, the above expression simplifies to

$$d\ln C_1 = \frac{1}{1+\theta\mu_1\mu_2}\frac{w_1}{P_1 C_1} d\ln w_1. \tag{E.8}$$

Note that the only source of variation in the above expression is $\frac{w_1}{P_1 C_1}$. Therefore, the elasticity of consumption with respect to wages is systematically higher for locations with a fiscal surplus ($w_1 - P_1 C_1 > 0$), and lower for locations with a fiscal deficit ($w_1 - P_1 C_1 < 0$). Finally, note that the population response in location 1 is given by

$$d\ln \mu_1 = \theta(1-\mu_1) d\ln C_1 \tag{E.9}$$

under log utility. Consequently,

$$d\ln \mu_1 = \theta(1-\mu_1)\frac{1}{1+\theta\mu_1\mu_2}\frac{w_1}{P_1 C_1} d\ln w_1. \tag{E.10}$$

Therefore, under our assumptions, the elasticity of population with respect to the local wage is strictly increasing in $\frac{w_1}{P_1 C_1}$ at the planner's solution.

F Quantitative Appendix

F.1 Data

We obtain state-level price indices from the Bureau of Economic Analysis (BEA). State-level per capita transfers and taxes are also from BEA. We obtain annual state-level bilateral migration flows in 2017 from the Census website,[F.1] where the underlying data are based on the American Community Survey. We translate into 5-year migration rates by raising to the power of five. We obtain state-level bilateral trade flows at the state level from the replication packages of Kleinman et al. (2023), which they construct using the Commodity Flow Survey. We further adjust these bilateral trade flows so that the total exports (including self-exports) equal state-level GDP in the

[F.1] https://www.census.gov/data/tables/time-series/demo/geographic-mobility/state-to-state-migration.html

BEA data.

F.2 Details on Calibration

Aggregate trade flows from location i to j, denoted as $x_{ij} \equiv w_i l_{ij}$, are given by

$$x_{ij} = \frac{(w_i/\mathcal{A}_{ij}(L_i))^{1-\sigma}}{\sum_k (w_k/\mathcal{A}_{kj}(L_k))^{1-\sigma}} \int P_j C_j(a) d\varphi_j(a). \tag{F.1}$$

Taking the ratio of x_{ij} to x_{jj}, we have

$$\frac{x_{ij}}{x_{jj}} = \frac{(w_i/\mathcal{A}_{ij}(L_i))^{1-\sigma}}{(w_j/\mathcal{A}_{jj}(L_j))^{1-\sigma}}, \tag{F.2}$$

which we can rewrite as

$$\frac{1}{\mathcal{A}_{ij}(L_i)} = \left(\frac{x_{ij}}{x_{jj}}\right)^{\frac{1}{1-\sigma}} \frac{w_j}{w_i} \frac{1}{\mathcal{A}_{jj}(L_j)}. \tag{F.3}$$

The price index of location j is

$$\begin{aligned} P_j &= \left[\sum_i (w_i/\mathcal{A}_{ij}(L_i))^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \\ &= \left[\sum_i \left(\left(\frac{x_{ij}}{x_{jj}}\right)^{\frac{1}{1-\sigma}} \frac{w_j}{w_i} \frac{1}{\mathcal{A}_{jj}(L_j)} w_i\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \\ &= \frac{w_j}{\mathcal{A}_{jj}(L_j)} \frac{1}{(x_{jj})^{\frac{1}{1-\sigma}}} \left[\sum_i x_{ij}\right]^{\frac{1}{1-\sigma}} \\ &= \frac{w_j}{A_{jj} L_j^{\alpha}} \frac{1}{(x_{jj})^{\frac{1}{1-\sigma}}} \left[\sum_i x_{ij}\right]^{\frac{1}{1-\sigma}}, \end{aligned} \tag{F.4}$$

where we used (F.3) in the second line. As a result, conditional on the choice of (σ, α), we can infer A_{jj} given data on trade flows $\{x_{ij}\}$, price indices $\{P_j\}$, population size L_j, and output per capita $w_j = \sum_j x_{ij}/L_i$:

$$A_{jj} = \frac{w_j}{P_j L_j^{\alpha}} \frac{1}{(x_{jj})^{\frac{1}{1-\sigma}}} \left[\sum_i x_{ij}\right]^{\frac{1}{1-\sigma}}. \tag{F.5}$$

With $\{A_{jj}\}$ in hand, we also infer all $\{A_{ij}\}$ using (F.3):

$$A_{ij} = \frac{1}{L_i^\alpha} \left(\frac{x_{jj}}{x_{ij}} \right)^{\frac{1}{1-\sigma}} \frac{w_i}{w_j} A_{jj} L_j^\alpha. \tag{F.6}$$

We choose the remaining parameter values $\{\{\chi_{ij}\}, \underline{a}, \beta\}$ by repeatedly solving the model to exactly match (i) migration flows in the data, (ii) the steady-state real interest rate target, and (iii) the target value of the marginal propensity to consume. We normalize $\chi_{ii} = 1$ for all i, since what matters for migration decisions is χ_{ij}/χ_{ii}. In calibrating migration costs, we use the following updating rule. Given the guess of $\{\chi_{ij}^{old}\}$, we can solve the model to obtain aggregate migration flows from region i to j:

$$\mu_{ij}^{model} \equiv \int \mu_{ij}(a) d\varphi_i. \tag{F.7}$$

Given the migration probabilities in the data, μ_{ij}^{data}, we update χ_{ij} as follows

$$\chi_{ij}^{new} = \xi \frac{\mu_{ij}^{data}}{\mu_{ij}^{model}} \chi_{ij}^{old} + (1-\xi)\chi_{ij}^{old}, \tag{F.8}$$

where $\xi \in (0, 1]$ is the degree of updating. For β and $\underline{a}$, we update using the bisection method.

F.3 Steady-State Computational Algorithms

We describe the following computational algorithm to solve the status quo and constrained-efficient allocations. For both cases, we describe the algorithm for the steady-state. The algorithm for the transitions are similar with everything indexed by time t.

F.3.1 Computational Algorithm for Status Quo Economy

In the steady-state, households solve

$$v_j(a) = \max_{C_j, \{\mu_{jk}\}, \tilde{a} \geq \underline{a}} u_j(C_j) + \beta\omega \left[\sum_k \mu_{jk} v_k(\tilde{a}) - \psi_j(\{\mu_{jk}\}) \right] \tag{F.9}$$

$$\text{s.t.} \quad P_j C_j + \tilde{a} = (1+r)a + w_j + T_j. \tag{F.10}$$

The first-order condition with respect to $\tilde{a}$ is

$$u_j'(C_j)/P_j \geq \beta\omega \left[\sum_j \mu_{jk} \partial_a v_k(\tilde{a}) \right] \tag{F.11}$$

with equality whenever $\tilde{a} > \underline{a}$.

For the inner problem, where we solve the Bellman equation given prices, we proceed as follows. The algorithm extends the endogenous gridpoint method of Carroll (2006) to incorporate dynamic discrete choices. Let $\mathbb{A} \equiv [a_1, \ldots, a_I]$ denote the gridpoints of assets.

1. For each gridpoint in $\tilde{a} \in \mathbb{A}$, guess $\{v_k(\tilde{a})\}$.

2. Given $\{v_k(\tilde{a})\}$, one can compute migration probabilities conditional on saving $\tilde{a}$ by solving

$$\{\mu_{jk}^{EGM}(\tilde{a})\} \in \arg\max_{\{\mu_{jk}\}} \left\{ \sum_k \mu_{jk} v_k(\tilde{a}) - \psi_{jt}(\{\mu_{jk}\}) \right\} \tag{F.12}$$

   for each j and $\tilde{a} \in \mathbb{A}$. Under a type-I extreme value distribution (55), this is analytical and immediate to obtain:

$$\mu_{jk}^{EGM}(\tilde{a}) = \frac{\chi_{jk} \exp(\theta v_k(\tilde{a}))}{\sum_l \chi_{jl} \exp(\theta v_l(\tilde{a}))}. \tag{F.13}$$

3. Assuming the first-order condition holds with equality, invert the consumption using

$$C_j^{EGM}(\tilde{a}) = u_j'^{-1} \left(P_j \beta \omega \left[\sum_j \mu_{jk}^{EGM}(\tilde{a}) \partial_a v_k(\tilde{a}) \right] \right) \tag{F.14}$$

   for each j and $\tilde{a} \in \mathbb{A}$. Then, we are able to obtain the current asset level that is consistent with next period savings $\tilde{a}$ and a non-binding borrowing constraint:

$$a_j^{EGM}(\tilde{a}) = \frac{1}{1+r} \left(\tilde{a} + P_j C_j^{EGM}(\tilde{a}) - w_j - T_j \right). \tag{F.15}$$

4. For a such that $a \leq a_j^{EGM}(\underline{a})$, the borrowing constraint is binding. Therefore, we recover the saving policy functions as follows:

$$\tilde{a}_j(a) = \begin{cases} a_j^{EGM,-1}(a) & \text{if } a > a_j^{EGM}(\underline{a}) \\ \underline{a} & \text{if } a \leq a_j^{EGM}(\underline{a}) \end{cases}, \tag{F.16}$$

   where $a_j^{EGM,-1}$ denotes the inverse function of $a_j^{EGM}(a)$.

5. The migration policies are given by

$$\mu_{jk}(a) = \mu_{jk}^{EGM}(\tilde{a}_j(a)), \tag{F.17}$$

and the consumption function is

$$C_j(a) = \frac{1}{P_j}\left((1+r)a + w_j - \tilde{a}_j(a)\right). \tag{F.18}$$

Now we can update the value function as

$$v_j^{new}(a) = u_j(C_j(a)) + \beta\omega\left[\sum_k \mu_{jk}(a) v_k(\tilde{a}_j(a)) - \psi_{jt}(\{\mu_{jk}(a)\})\right]. \tag{F.19}$$

If $|v_j^{new}(a) - v_j(a)| < tol$, we are done. Otherwise, go back to 2 with $v_j(a) := v_j^{new}(a)$.

The outer problem iterates over prices $\{r, \{w_{j,}, L_j\}\}$. We divide the outer problem into two layers. In the inner layer, we iterate over r to clear the bond market. In the outer layer, we iterate $\{w_j, L_j\}$ to clear the final goods market for each location.

1. Guess $\{w_j, L_j\}$, where we take location 1's wage as the numeraire: $w_1 = 1$.

2. Given $\{w_i, L_i\}$, compute the price indices in each location:

$$P_j = \left[\sum_i \left(\frac{w_i}{A_{ij}\, L_i^{\alpha}}\right)^{1-\sigma}\right]^{1/(1-\sigma)}. \tag{F.20}$$

3. Given $\{w_j\}$, iterate over r or β until the bond market clears: $\sum_j \int a d\varphi_j(a) \approx 0$. We use bisection to update r or β. We iterate over r when we solve for counterfactuals. We iterate over β when we match the target interest rate r.

4. We then update wages $\{w_j\}$ and population size $\{L_j\}$ as follows. Given the implied distribution φ_j and consumption policy functions $C_j(a)$ from the guess of $\{w_j\}$, we compute

$$w_i^{new} = \xi^w\left[\frac{1}{L_i}\sum_j \frac{(1/(A_{ij}L_i^{\alpha}))^{1-\sigma}}{\sum(w_l/(A_{lj}L_l^{\alpha}))^{1-\sigma}}\int P_j C_j(a) d\varphi_j(a)\right]^{\frac{1}{\sigma}} + (1-\xi^w)w_i \tag{F.21}$$

$$L_i^{new} = \xi^L \int d\varphi_i(a) + (1-\xi^L)L_i \tag{F.22}$$

where $\xi \in (0,1]$ is the degree of updating. If $|w_i^{new} - w_i| < tol$ and $|L_i^{new} - L_i| < tol$ for all i, we are done. Otherwise, set $w_i := w_i^{new}$ and $L_i := L_i^{new}$ and go back to step 1.

Practically, the above algorithm finds the equilibrium prices orders of magnitude faster than more conventional algorithms such as Newton's method.

F.3.2 Computational Algorithm for the Constrained-Efficient Allocation

Throughout, we impose the parametric functional form (55) that we use in the quantitative exercise. We first derive the optimality conditions under this function form, which will be useful for our computation. We then explain how we can efficiently solve the constrained-efficient allocation on the computer.

The problem in the steady-state is

$$S_i(v) = \max_{C_i, \{\tilde{v}_{ij}, \mu_{ij}\}} w_i\left(1+\alpha_i\right) - P_i C_i + (1-\omega) S_i^n + \frac{1}{R}\omega \sum_j \mu_{ij} S_j(\tilde{v}_{ij}) \tag{F.23}$$

$$\text{s.t.} \quad v = u_i(C_i) + \beta\omega \left[\sum_j \mu_{ij}\tilde{v}_{ij} - \psi_i(\{\mu_{ij}\})\right] \tag{F.24}$$

$$\{\mu_{ij}\} \in \arg\max_{\{\tilde{\mu}_{ij}\}} \left\{\sum_j \tilde{\mu}_{ij}\tilde{v}_{ij} - \psi_i(\{\tilde{\mu}_{ij}\})\right\} \tag{F.25}$$

and the value of the newborn in the steady-state is

$$S_i^n = \max_{v_i^n, \{\tilde{v}_{ij}^n, \mu_{ij}^n\}} \Lambda_i v_{it}^n + \frac{1}{R}\sum_j \mu_{ij}^n S_j(\tilde{v}_{ij}^n) \tag{F.26}$$

$$\text{s.t.} \qquad v_{it}^n = \beta \sum_j \left[\mu_{ij}\tilde{v}_{ij}^n - \psi_i(\{\mu_{ij}^n\})\right] \tag{F.27}$$

$$\{\mu_{ij}^n\} \in \arg\max_{\{\tilde{\mu}_{ij}^n\}} \left\{\sum_j \tilde{\mu}_{ij}^n \tilde{v}_{ij}^n - \psi_i(\{\tilde{\mu}_{ij}^n\})\right\}, \tag{F.28}$$

A challenge in numerically solving the above problem is the dimensionality of the control variables. We need to optimize over continuation value for each location. However, we show below that, with our functional form assumption (55), the problem essentially collapses to a one-dimensional optimization problem.

We first describe the problem for the incumbent generation (F.23). We solve for C_j to rewrite the problem as

$$S_i(v) = \max_{C_i, \{\tilde{v}_{ij}, \mu_{ij}\}} w_i\left(1+\alpha_i\right) - P_i C_i + (1-\omega)\Lambda_i S_i^n + \frac{1}{R}\omega \sum_j \mu_{ij} S_j(\tilde{v}_{ij}) \tag{F.29}$$

$$\text{s.t.} \quad C_i = u_i^{-1}\left(v - \beta\omega\left[\sum_j \mu_{ij}\tilde{v}_{ij} - \psi_i(\{\mu_{ij}\})\right]\right) \tag{F.30}$$

$$\{\mu_{ij}\} \in \arg\max_{\{\tilde{\mu}_{ij}\}} \left\{\sum_j \tilde{\mu}_{ij}\tilde{v}_{ij} - \psi_i(\{\tilde{\mu}_{ij}\})\right\} \tag{F.31}$$

The first-order conditions with respect to v_{im} are

$$\frac{\omega}{R}\mu_{imt}S'_m(\tilde{v}_{im}) + \frac{P_j}{u'_i(C_i)}\beta\omega\mu_{imt} + \frac{\omega}{R}\sum_k \frac{\partial\mu_{ik}}{\partial v_{im}} S_k(\tilde{v}_{ik}) = 0, \tag{F.32}$$

where we have omitted the dependence for brevity. Under the logit specification (55), we have

$$\frac{\partial\mu_{jk}}{\partial\tilde{v}_m} = \begin{cases} \theta\mu_{jk}\left(1-\mu_{jk}\right) & \text{for } k = m \\ -\theta\mu_{jk}\mu_{jm} & \text{for } k \neq m \end{cases}. \tag{F.33}$$

Therefore, the FOC simplifies to

$$S'_m(\tilde{v}_{im}) + \beta R\frac{P_i}{u'_i(C_i)} + \theta S_m(\tilde{v}_{im}) - \theta\sum_k \mu_{ik}S_k(\tilde{v}_{ik}) = 0. \tag{F.34}$$

From this expression, the policy functions must satisfy:

$$S'_m(\tilde{v}_{im}(v)) + \theta S_m(\tilde{v}_{im}(v)) = S'_n(\tilde{v}_{in}(v)) + \theta S_n(\tilde{v}_{in}(n)) \tag{F.35}$$

$$= \theta\sum_k \mu_{ik}S_k(\tilde{v}_{ik}(v)) - \beta R\frac{P_i}{u'_i(C_i(v))} \tag{F.36}$$

$$\equiv M_i(v) \tag{F.37}$$

for all m and n. This observation leads to a substantial simplification. Instead of optimizing over $\{v_{im}(v)\}_m$, we instead optimize over a one-dimensional object $M_i(v)$. Given the guess of $M_i(v)$, we can immediately obtain $v_{im}(v)$ by solving

$$S'_m(\tilde{v}_{im}(v)) + \theta S_m(\tilde{v}_{im}(v)) = M_i(v) \tag{F.38}$$

for each m. Once we obtain $\{v_{im}(v)\}$, we can obtain $\{\mu_{imt}(v)\}$ using (F.31). With $\{v_{im}(v)\}$ and $\{\mu_{imt}(v)\}$ in hand, consumption is residually determined from the promise-keeping constraint (F.30):

$$C_i(v) = u_i^{-1}\left(v - \beta\omega\left[\sum_j \mu_{ij}(v)\tilde{v}_{ij}(v) - \psi_i(\{\mu_{ij}(v)\}_j)\right]\right). \tag{F.39}$$

Given all the steps for a given guess of $M_i(v)$, we can search for the optimal $M_i(v)$ using a standard one-dimensional optimization routine such as the Brent method. In practice, we can obtain further speed gains with the endogenous gridpoint method by Carroll (2006). Below, we describe the algorithm for value function iteration that relies on the endogenous gridpoint method.

The newborn's problem (F.26) can be solved similarly, or is even more simplly. The first-order condition with respect to v^n_{im} is

$$S'_m(\tilde{v}^n_{im}) + \beta R \Lambda_i + \theta S_m(\tilde{v}^n_{im}) - \theta \sum_k \mu^n_{ik} S_k(\tilde{v}^n_{ik}) = 0, \tag{F.40}$$

which is analogous to (F.34). Therefore, it must be that

$$S'_m(\tilde{v}^n_{im}) + \theta S_m(\tilde{v}^n_{im}) = M^n_i \tag{F.41}$$

for some M^n_i. For a given guess of M^n_i, we can find the continuation value $\tilde{v}^n_{im}$ that is consistent with M^n_i for all m by inverting (F.41). Given $\tilde{v}^n_{im}$, we can find the migration probabilities using the incentive compatibility constraint (F.28).

Algorithm. We first describe the algorithm for solving the Bellman equation for a given vector of $\{w_i, P_i\}_i$. Note that with our functional form assumption (53), α_i is exogenously fixed at α. The outer loop updates $\{w_i, P_i\}$, which we describe later. We let $\mathbb{V} \equiv [v_1, \ldots, v_{N_V}]$ denote the gridpoint of promised utility and $\mathbb{M}$ denote the gridpoints for $M_{it}(v) \in \mathbb{M} \equiv [M_1, M_2, \ldots, M_{N_M}]$.

1. Guess the value function $S_{it}(v)$.
2. For each $i = 1, \ldots, J$,
   (a) For each $M \in \mathbb{M}$
      i. Compute $\tilde{v}^{EGM}_{im}(M)$ that is consistent with (F.38) where $M_{it}(v) = M$:

      $$S'_m(\tilde{v}^{EGM}_{im}(M)) + \theta S_m(\tilde{v}^{EGM}_{im}(M)) = M, \tag{F.42}$$

      for each $m = 1, \ldots, J$.
      ii. Using $\{\tilde{v}^{EGM}_{im}(M)\}$, compute $\{\mu^{EGM}_{im}(M)\}$ using (F.31) associated with $v_{im} = \tilde{v}^{EGM}_{im}(M)$.
      iii. Find $C_i(M)$ that is consistent with the optimality conditions (F.36) and (F.37):

      $$C^{EGM}_i(M) = u'^{-1}_i\left(\frac{\beta R P_i}{[\theta \sum_k \mu^{EGM}_{ik}(M) S_k(\tilde{v}^{EGM}_{ik}(M)) - M]}\right). \tag{F.43}$$

      iv. Now we can find the value of today's promised utility v that is consistent with

M using the promise keeping constraint (F.30):

$$v_i^{EGM}(M) = u_i(C_i^{EGM}(M)) + \beta\omega \left[\sum_j \mu_{ij}^{EGM}(M)\tilde{v}_{ij}^{EGM}(M) - \psi_i(\{\mu_{ij}^{EGM}(M)\}_j) \right]. \tag{F.44}$$

(b) Now we invert the mapping of $v_i^{EGM}(M)$ to obtain the optimal M for each $v \in \mathbb{V}$: $M_i(v) \equiv v_{it}^{EGM,-1}(M)$. With $M_i(v)$ for each $v \in \mathbb{V}$ in hand, we can compute all the associate policy functions from the previous step.

(c) For the newborn's problem (F.26), we simply optimize over M_i^n and finds associated continuation values $\{v_{im}^n\}_m$ using (F.41) to maximize the right hand side of (F.26) to obtain S_{it}^n.

(d) Now we can update the value function:

$$S_i^{new}(v) = w_i(1+\alpha_i) - P_i C_i(v) + (1-\omega)\Lambda_i S_i^n + \frac{1}{R}\omega \sum_j \mu_{ij}(v) S_j(\tilde{v}_{ij}(v)). \tag{F.45}$$

3. If $|S_{it}(v)^{new} - S_{it}(v)| < tol$ for all i and $v \in \mathbb{V}$, the value function has converged. If not, update the value function, $S_{it}(v) := S_{it}(v)^{new}$ and go back to step 2.

The outer loop looks for the Lagrange multipliers $\{w_i, P_i\}$ that are consistent with the resource constraints. In practice, it is slightly easier to iterate over $\{w_i, L_i\}$ instead of $\{w_i, P_i\}$, although they are conceptually equivalent. We proceed as follows.

1. Guess $\{w_i, L_i\}$.

2. Given $\{w_i, L_i\}$, compute the implied $\{P_j\}$ with the CES price index:

$$P_j = \left[\sum_i \left(\frac{w_i}{A_{ij}\, L_i^{\alpha}} \right)^{1-\sigma} \right]^{1/(1-\sigma)}. \tag{F.46}$$

3. With $\{w_i, P_i\}$ in hand, solve the Bellman equation using the algorithm described above. This gives us the consumption policy functions $\{C_j(v)\}$ and the steady-state distribution $\{\phi_j\}$.

4. We then update wages $\{w_j\}$ and population size $\{L_j\}$ as follows. Given the distribution ϕ_j

and consumption policy functions $C_j(a)$ from the guess $\{w_j\}$, we compute

$$w_i^{new} = \xi^w \left[\frac{1}{L_i} \sum_j \frac{(1/(A_{ij} L_i^\alpha))^{1-\sigma}}{\sum (w_l/(A_{lj} L_l^\alpha))^{1-\sigma}} \int P_j C_j(v) d\phi_j \right]^{\frac{1}{\sigma}} + (1-\xi^w) w_i \tag{F.47}$$

$$L_i^{new} = \xi^L \int d\varphi_i + (1-\xi^L) L_i \tag{F.48}$$

where $\xi^w, \xi^L \in (0,1]$ are the degrees of updating. If $|w_i^{new} - w_i| < tol$ and $|L_i^{new} - L_i| < tol$ for all i, we are done. Otherwise, set $w_i := w_i^{new}$ and $L_i := L_i^{new}$ and go back to step 2.

F.4 Constrained-Efficient Allocation with Aggregate Shocks

For notational simplicity, we drop x superscript and let t denote the time elapsed since the arrival of the aggregate shock. Note that with our assumption that $p \to 0$, the economy is in a deterministic steady-state before the arrival of the shock. For this reason, we drop the time subscript for variables before the arrival of the shock. We will describe the environment under the specific functional form assumptions of our calibration. The general case is similar, albeit the notation is slightly more complex.

The first-order condition with respect to v_{ij0} is

$$p \frac{1}{R} \mu_{ij0} S'_{j0}(v_{ij0}) + p \frac{1}{R} \sum_k \frac{\partial \mu_{ik0}}{\partial v_{ij0}} S_{k0}(v_{ik1}) + p\beta \frac{P_i}{u'_i(C_i(v))} \mu_{ij0} = 0. \tag{F.49}$$

Notice that $P_i / u'_i(C_i(v))$ is evaluated at the initial steady-state. Condition (F.49) and associated policy functions $\{v_{ij0}(v), \mu_{ij0}(v)\}$, along with the steady-state distribution, determine the initial distribution upon the arrival of the aggregate shock: $\{\phi_{i0}(v)\}$. We denote it $\boldsymbol{\phi}_0$ in a (discretized) vector format.

After the arrival of the aggregate shock, given the induced changes in the sequence of the aggregates, the component planning problem solves (31) and (34). In general, the sequence of aggregates consists of $\{P_{jt}, w_{jt}, \alpha_{jt}, f_{jt}\}_{t=0}^\infty$. Under our calibration, it suffices to track $\{w_{jt}, L_{jt}, \{A_{jkt}\}\}_{t=0}^\infty$. This is because, given $\{w_{jt}, L_{jt}, \{A_{jkt}\}\}_{t=0}^\infty$, one can immediately compute P_{it} using (F.46). For this reason, we treat $\{w_{jt}, L_{jt}, \{A_{jkt}\}\}_{t=0}^\infty$ as the sequence of the aggregates that we need to keep track of. We denote them $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})$ in a vector format (stacking both j and t).

Let $\{C_{it}(v; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}), \{\mu_{ijt}(v; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}), v_{ijt+1}(v; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})\}\}$ be the policy functions associated with (31) and (34) as well as a sequence of $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})$. Given the policy functions and the initial distribution $\boldsymbol{\phi}_0$, we can compute the distribution at any point in time, which we denote as $\boldsymbol{\phi}_t(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0)$. The aggregate sequence in turn needs to be consistent with the market clearing condition and the consistency in the population distribution:

$$
\begin{aligned}
0 &= w_{it}L_{it} - \frac{(w_{it}/(A_{ijt}L_i^{\alpha}))^{1-\sigma}}{\sum_l (w_{lt}/(A_{ljt}L_l^{\alpha}))^{1-\sigma}} \int P_{jt}C_{jt}(v; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}) d\phi_{jt}(v; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0) \\
&\equiv \mathcal{F}_t^C(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0) \qquad \text{(F.50)} \\
0 &= L_{it} - \int d\phi_{it}(v; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0) \\
&\equiv \mathcal{F}_t^L(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0). \qquad \text{(F.51)}
\end{aligned}
$$

Note that, from (F.49), the initial distribution $\boldsymbol{\phi}_0$ is solely determined by the sequence of aggregates $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})$ because S_{j0} is determined by (31) and (34) given $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})$. We write this relationship as follows:

$$\boldsymbol{\phi}_0 = \Phi_0(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}). \qquad \text{(F.52)}$$

Consequently, obtaining the transition dynamics in response to technology shocks $\boldsymbol{A}$, amounts to solving the following system of equations with respect to $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{\phi}_0)$:

$$
\begin{aligned}
0 &= \mathcal{F}_t^C(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0) \qquad \text{(F.53)} \\
0 &= \mathcal{F}_t^L(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0) \qquad \text{(F.54)} \\
\boldsymbol{\phi}_0 &= \Phi_0(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}). \qquad \text{(F.55)}
\end{aligned}
$$

We look for first-order solutions with respect to the aggregate shocks $\boldsymbol{A}$. For this goal, the Jacobian of $\mathcal{F}_t^C$, $\mathcal{F}_t^L$, and Φ_0 with respect to $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0)$ are sufficient. The Jacobians of $\mathcal{F}_t^C$ and $\mathcal{F}_t^L$ correspond to the Sequence Space Jacobian (Auclert et al. 2021), and we can efficiently compute them using the single backward iteration and the fake news algorithms, as explained by Auclert et al. (2021). The Jacobian of Φ_0 can easily be obtained during the process of the single backward iteration. The algorithm here resembles that of Auclert et al. (2024) in the context of the endogenous portfolio choice in heterogeneous agent models. In implementing it, we truncate the transitions to a horizon of 500 years.

F.5 Status Quo with Aggregate Shocks

We now describe the status quo economy with aggregate shocks. We maintain the assumptions described in Section 7. Under the assumption that $p \to 0$, the economy is in a deterministic steady-state before the arrival of the aggregate shock. For notational simplicity, we drop x superscript and let t denote the time elapsed since the arrival of the aggregate shock. We drop time subscript for variables before the arrival of the shock. As in Appendix F.4, we describe the environment under the specific functional form assumptions of our calibration.

Before the arrival of the aggregate shock, the Bellman equation of a household living in location j with asset holding a is

$$v_i(a) = \max_{C_i, a', \{\mu_{ij}, \mu_{ij0}\}} u_i(C_i) + (1-p)\beta\omega \left[\sum_j \mu_{ij} v_j(a') - \psi_i(\{\mu_{ij}\})\right] + p\beta\omega \left[\sum_j \mu_{ij0} v_{j0}(a') - \psi_{it}(\{\mu_{ij0}\})\right] \tag{F.56}$$

$$\text{s.t.} \quad P_i C_i + a' = (1+r)a + w_i + T_i \tag{F.57}$$

$$a' \geq \underline{a}. \tag{F.58}$$

Here, our timing assumption implies that the saving policy function $a'_i(a)$ is not contingent on the arrival of the aggregate shock, but the migration probabilities $\{\mu_{ij}(a), \mu_{ij0}(a)\}$ are. Because of this, the initial distribution immediately after the arrival of the aggregate shock, which we denote $\{\varphi_0(a)\}$, is generally different from the steady-state distribution. We denote the initial distribution $\boldsymbol{\varphi}_0$ in a (discretized) vector format.

After the arrival of the aggregate shock, which induces a sequence of aggregates, the Bellman equations are given by (42). As in Appendix F.4, the sequence that matters for the Bellman equation is summarized by $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})$. We let $\{C_{it}(a; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}), a_{it+1}(a; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}), \{\mu_{ijt}(a; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})\}\}$ denote the policy functions following the arrival of the aggregate shocks. Given the policy functions and initial distribution $\boldsymbol{\varphi}_0$, we can compute the distribution at any point in time. We let $\{\varphi_t(a; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\varphi}_0)\}$ denote the distribution at time t as a function of $\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}$, and $\boldsymbol{\varphi}_0$.

The sequence of aggregates needs to satisfy the market clearing conditions as well as consistency in the population distribution:

$$0 = w_{it} L_{it} - \frac{(w_{it}/(A_{ijt} L_i^\alpha))^{1-\sigma}}{\sum_l (w_{lt}/(A_{ljt} L_l^\alpha))^{1-\sigma}} \int P_{jt} C_{jt}(a; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}) d\varphi_{jt}(a; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\varphi}_0) \equiv \mathcal{F}_t^C(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\varphi}_0) \tag{F.59}$$

$$0 = L_{it} - \int d\varphi_{it}(a; \boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\varphi}_0) \equiv \mathcal{F}_t^L(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\varphi}_0). \tag{F.60}$$

Note that from (F.56), the initial migration probabilities $\{\mu_{ij0}(a)\}$ are determined by the initial value functions $\{v_{j0}(a)\}$, which, in turn, are determined by the aggregate sequence $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})$ from (42). This implies that the initial distribution $\boldsymbol{\varphi}_0$ is a function of $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A})$. We write this relationship as

$$\boldsymbol{\varphi}_0 = \Phi_0(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}). \tag{F.61}$$

Consequently, solving the transition dynamics following the aggregate shock amounts to finding the $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{\varphi}_0)$ that satisfies the following system of equations:

$$0 = \mathcal{F}_t^C(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\varphi}_0) \tag{F.62}$$

$$0 = \mathcal{F}_t^L(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\varphi}_0) \tag{F.63}$$

$$\boldsymbol{\varphi}_0 = \Phi_0(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}). \tag{F.64}$$

As in Appendix F.4, the Jacobian of $\mathcal{F}_t^C$, $\mathcal{F}_t^L$, and Φ_0 with respect to $(\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{A}, \boldsymbol{\phi}_0)$ are sufficient to compute the first-order solutions. The Jacobian of $\mathcal{F}_t^C$ and $\mathcal{F}_t^L$ correspond to the Sequence Space Jacobian (Auclert et al. 2021). The Jacobian of Φ_0 can easily be obtained during the process of the single backward iteration. Again, the algorithm here is similar to that in Auclert et al. (2024). In implementing it, we truncate the transitions to a horizon of 500 years.

F.6 Decomposing Welfare Gain

The welfare of a newborn in the steady-state is given by

$$W^{newborn} = \sum_i \Lambda_i (1-\omega) L_{it} v_{it}^n, \tag{F.65}$$

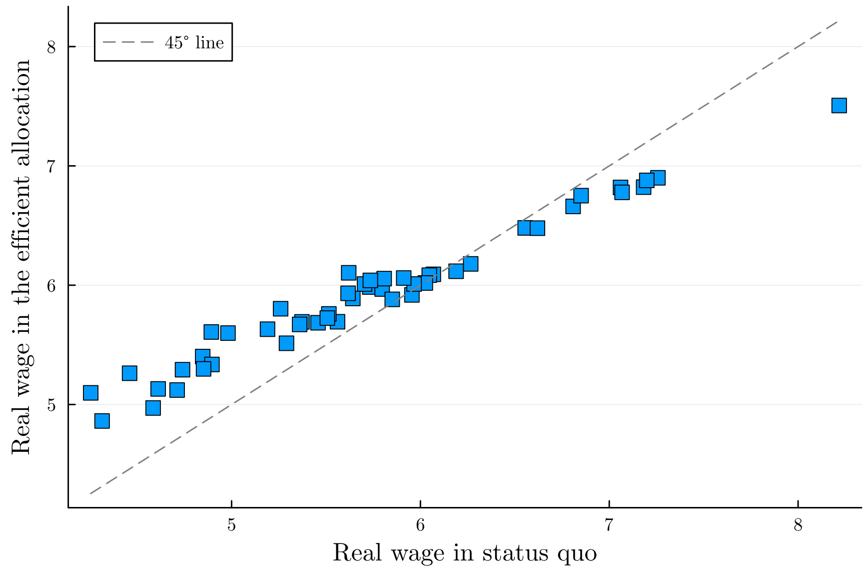
where L_{it} is the population size of location i, and v_{it}^n is the utilitarian welfare of newborns born in location i. Suppose the welfare of the newborn in the old economy is $W^{newborn}$, but it is $W^{newborn\prime}$ in the new economy. Likewise, let all the variables with and without prime denote those in the new and old economy, respectively. We can express the changes in welfare as

$$W^{newborn\prime} - W^{newborn} = \underbrace{\sum_i \Lambda_i (1-\omega) v_{it}^n \left[L_{it}' - L_{it}\right]}_{\text{(i) between location}} + \underbrace{\sum_i \Lambda_i (1-\omega) L_{it} \left[v_{it}^{n\prime} - v_{it}^n\right]}_{\text{(ii) within location}}. \tag{F.66}$$

We apply this decomposition to the welfare change from the status-quo economy to the constrained-efficient allocation. Expressed as a fraction of the consumption equivalent welfare gain, 50.2% is attributed to the (i) between location component and 49.8% is attributed to the (ii) between location component.

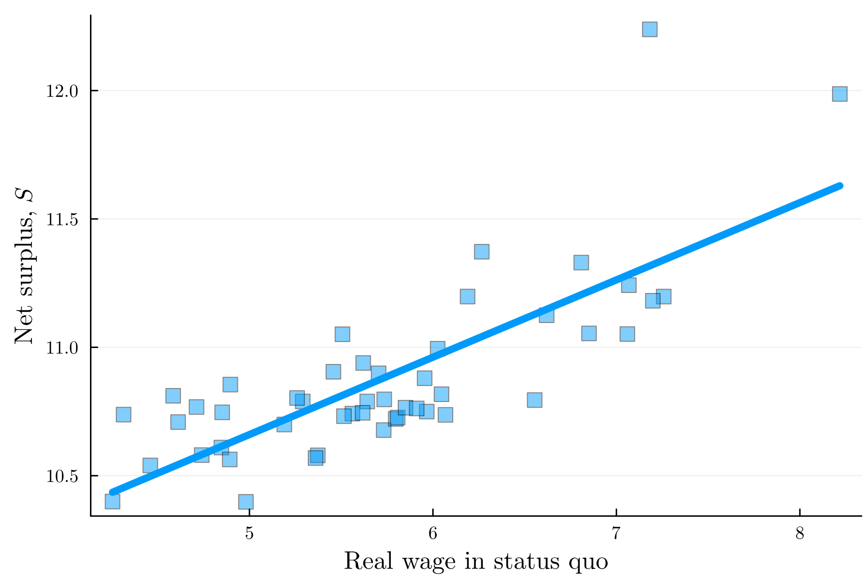
F.7 Additional Figures and Tables

Figure F.1: Real Wage: Status Quo vs. Planner



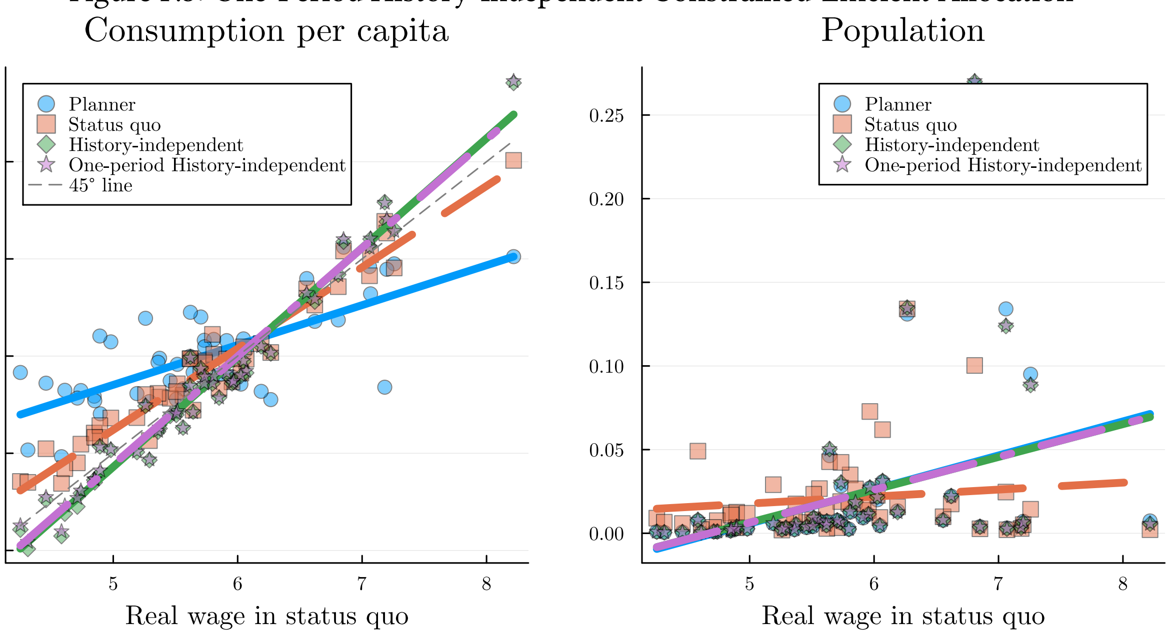
Note: The figure compares the real wage w_j/P_j in the constrained-efficient allocation (y axis) to that in the status-quo economy (x axis). Each square corresponds to a US state, and the dashed grey line is a 45-degree line.

Figure F.2: Net Surplus S in the Planner's Solution and Real Wage Status Quo



Note: The figure compares the population weighted average of net surplus S_j in the planner's solution (y axis) to the real wage in the status quo economy (x axis). Each square corresponds to a US state, and the solid blue line is the best linear fit.

Figure F.3: One-Period History-Independent Constrained-Efficient Allocation



Note: The figure presents a version of Figure 3, where we also plot purple stars for the "one-period history-independent constrained-efficient allocation" as described in Appendix C.