From Savings Push to Demand Pull: Dynamic Consumption Externalities and Optimal Industrial Upgrading*

Xuewen Liu† Sichuang Xu‡

First version: September 2022; This version: October 2025

Abstract

We develop a tractable macro-development model to study industrial upgrading in catching-up economies. The model, incorporating hierarchical demand in preferences and learning-by-doing in production into a two-sector growth framework, captures a dynamic feedback loop: income growth expands consumption frontiers (income-push), while demand for advanced goods drives technological progress (demand-pull). The laissez-faire equilibrium exhibits two-sided inefficiencies due to dynamic production externalities—excessive early-stage consumption delays capital accumulation, whereas insufficient later-stage consumption limits productivity spillovers. Optimal policy entails a two-phase intervention: initially suppressing consumption to prioritize investment, followed by stimulating consumption to exploit spillovers. Targeted industrial policies facilitate the timely rebalancing toward consumption-led growth, averting productivity traps. We further provide empirical evidence and quantify the model using cross-country data and Chinese micro data.

Keywords: Hierarchical demand, learning-by-doing, dynamic externalities, two-phase policy

*For helpful comments, we thank Linyi Cao, Ding Dong, Xiang Fang, Yan Ji, Ernest Liu, Chang Ma, Thomas Sargent, Yongseok Shin, Dan Su, Pengfei Wang, Yong Wang, Yi Wen, Danyang Xie, Bo Zhao, Lijun Zhu, Xiaodong Zhu, and seminar participants at 2025 World Congress of the Econometric Society, 2025 ICCDS, 2023 CICM, HKU-UCL-ESRC summer workshop, 2023 PHBS Workshop in Macroeconomics and Finance, 2023 International Workshop on New Structural Economics, 2023 NCER-CCER Workshop, CKGSB, HKU, HKUST, Fudan University, Shanghai Jiaotong University, and Peking University (INSE and GSM). The paper was previously circulated under the titles "Consumption-Led Industrial Upgrading" and "Dynamic Big Push".

†HKU Business School, University of Hong Kong (HKU). Email: xuewenl@hku.hk.

‡School of Management and Economics, Chinese University of Hong Kong (Shenzhen). Email: xusichuang@cuhk.edu.cn.

As living standards rise and consumption upgrades, consumer demand for freshness and timeliness surges. A wide range of products—fruits, vegetables, meats, medicines—requires precise temperatures and conditions... Yet, China's cold-chain industry is still in its infancy and has significant room for efficiency gains. Currently, wastage rates are as high as 15-20%, compared to 5% in Europe and the U.S., and just 2% in Japan... In fact, the more consumers demand, the faster efficiency improves.

— quoted from a news article entitled "China's Economic Transformation and Upgrading: A View from the Cold-Chain Industry," published in *Caixin* on June 29, 2020.[1]

1 Introduction

In search of the engine of sustained growth, a vast literature studies production externalities and spillovers (learning by doing) as a source of endogenous growth. The pioneering work of Arrow (1962), Romer (1986), and Lucas (1988), along with follow-up work, models externalities and spillovers in physical (human) capital. The policy implication of these models is that, in competitive equilibrium, consumption is too high and investment is too low at any given point in time. Some later work in this literature, notably Stokey (1988) and Matsuyama (2002), instead emphasizes that production externalities arise from the introduction of new consumption goods.[2] However, because of the endowment-economy setting with only labor and no capital, these latter models have little to say about transition and growth and, in particular, are silent on policy implications for development paths involving capital accumulation, without emphasizing the differences between the social planner's solution and the competitive equilibrium. Our paper aims to advance this literature by studying externalities and spillovers in the consumption goods sector, as in Stokey (1988) and Matsuyama (2002), while allowing for endogenous capital accumulation.

Building a formal and tractable model to study our research question is challenging. First, the model framework needs to accommodate an unbounded number of potential new consumer goods to be introduced, with demand for sequentially introduced new goods feeding back into the production side. Second, the externality that arises from consumption is inherently a dynamic and *recursive* problem (i.e., externalities induced by consumption generates new externalities), making its characterization difficult. Our paper overcomes these hurdles and develops a tractable model that features a standard growth framework with endogenous capital accumulation, allowing us to study intertemporal savings decisions and transitional dynamics, which are particularly relevant for developing countries, and to derive optimal development policies.

[1] Caixin, a leading newspaper in China. https://tangya.blog.caixin.com/archives/230802 (available in Chinese).

[2] Matsuyama (2002) focuses on industry-specific learning-by-doing, while Stokey (1988) focuses on inter-industry spillovers of learning-by-doing.

Specifically, we build a tractable macro-development growth model that incorporates hierarchical-demand preferences on the demand side and learning-by-doing on the production side. The model captures a growth pattern of "income-push new markets" and "demand-pull technological progress", i.e., income growth expands consumption frontiers (income-push), while demand for advanced goods drives technological progress (demand-pull). This gives rise to an endogenous dual-sector economic structure—the coexistence of technologically-mature industries and technologically-immature industries. The dynamically evolving dual-sector economic structure helps to shed light on industrial upgrading in catching-up economies.

In particular, the model analytically characterizes consumption-induced *dynamic* production externalities along the development path, yielding policy implications. The dynamic externalities render laissez-faire two-sidedly inefficient: excessive early-stage consumption delays capital accumulation, while insufficient later-stage consumption limits productivity spillovers. The optimal policy involves a two-phase intervention: initially suppressing consumption to prioritize investment, followed by stimulating consumption to exploit spillovers. Our finding that early-stage capital accumulation is a pathway to long-run sustainable growth speaks to the development experience of East Asian economies, characterized by rapid capital accumulation, and perhaps also to Latin America, often marked by conspicuous consumption and debt traps.[3] Equally importantly, our finding that the endogenous "sweet spot" in the development path—where the economy rebalances from rapid capital accumulation to consumption-led growth—occurs too late under laissez-faire suggests the need for targeted industrial policies, such as subsidies for emerging consumer-durables industries,[4] to facilitate timely rebalancing in emerging economies.

We present a two-sector AK growth model à la Rebelo (1991), with one sector using capital to produce consumption goods and the other using capital to produce investment goods (new capital). Within the consumption goods sector, we introduce a rich structure: households consume a consumption bundle comprising a continuum of indexed varieties, with hierarchical preferences over varieties as in Foellmi and Zweimüller (2006, 2008)—that is, a higher-indexed variety yields lower marginal utility. On the production side, the consumption-goods sector is endogenously split into two subsets of industries. One subset, producing relatively basic varieties, operates with *mature technology* at reduced marginal cost; the remaining subset, producing relatively advanced varieties, operates with *immature technology* at full marginal cost. The boundary variety that separates these two subsets of industries is called the "aggregate technology frontier". The "consumption frontier" denotes the cutoff variety consumed in equilibrium—the highest indexed variety affordable given income and relative prices.

[3] Brunnermeier, Gourinchas, and Itskhoki (2018) suggest that over-consumption with foreign borrowing in the early stages can lead to a sudden stop in the long run.

[4] See, e.g., the evidence in Akcigit et al. (2025) and Barwick et al. (2025).

In a dynamic economy, even without capital accumulation, a dynamic feedback loop arises between two frontiers, similar in spirit to Stokey (1988) and Matsuyama (2002). Specifically, the consumption frontier is pinned down by the income level and the aggregate technology frontier, with the former increasing along with the latter. This is because as the productivity of primitive varieties improve and their prices fall, the income effect dominates the substitution effect (akin to Engel's law), shifting demand toward more advanced varieties. In turn, as the consumption frontier expands, learning by doing in the production of newly demanded varieties accumulates technical know-how ("experience capital"), pushing the technology frontier forward. Over time, therefore, the two frontiers advance alternately through the forces of "income-push new markets" and "demand-pull technological progress".

In our dynamic economy with endogenous capital accumulation, the central decision is intertemporal: how much capital to allocate to producing consumption goods versus investment goods for future capital. Importantly, consumption generates production externalities: a larger market size induced by more consumption today reduces future marginal costs of production through learning-by-doing, creating spillovers. The planner internalizes this additional effect of consumption, whereas individual households do not in competitive equilibrium. The externality the planner considers is of a dynamic nature: more consumption today triggers spillovers starting tomorrow, while deferring consumption via investment shifts those spillovers to the day after tomorrow. That is, the production externality associated with consumption involves an *intertemporal tradeoff*.

Our primary finding concerns the optimal development policy. The laissez-faire equilibrium is two-sidedly inefficient due to dynamic externalities, characterized by early-stage over-consumption and later-stage under-consumption relative to the planner's allocation. This calls for a two-phase, two-sided intervention: initially suppressing consumption, then later stimulating it.

The source of this two-sided inefficiency is the time-varying value of consumption-induced externalities. Early in development, externalities are small for two reasons. First, consumption concentrates on basic goods whose production generates limited learning-by-doing spillovers compared to advanced goods that dominate later stages ("asymmetric externalities" over time). Second, when capital is scarce and baseline technology is relatively abundant, capital-technology complementarity implies that the marginal value of incremental technological improvements is low—additional innovations yield limited productivity gains without sufficient capital to utilize them. In this phase, deferring consumption to build capital maximizes future externalities. As development proceeds, capital abundance and a shift toward advanced goods increase the value of externalities: learning spillovers grow larger, and capital-technology complementarity raises the marginal returns to technological improvement. At this stage, stimulating consumption best exploits productivity spillovers. The planner thus transitions from suppression to stimulation, engineering a "demand pull" to accelerate industrialization. Targeted industrial policy can facilitate this transition by subsidizing

"emerging industries"—those that are neither too mature nor too premature.

The transition path also shapes long-run growth, implying the possibility of a development trap. The right policy affects not only the dynamics of transition but also the steady-state growth rate. In our two-sector economy, technological progress in the consumption-goods sector spills over to the investment-goods sector. If the consumption-goods sector cannot grow fast enough during transition, the investment-goods sector may fail to upgrade, becoming permanently stuck at a lower productivity level, which would result in a lower steady-state growth rate.

Extending the model to an open-economy setting strengthens the mechanisms under plausible conditions. Trade-driven capital inflows accelerate industrial upgrading, reducing production costs and expanding the range of exportable consumption goods. This creates a self-reinforcing cycle: cheaper, more competitive exports attract further capital inflows, which disproportionately boost domestic production and technological advancement. This amplification is manifested by the effect that a small expansion in foreign market size can induce a large increase in domestic consumption and technology levels.

To empirically validate our model's core mechanisms, we follow the empirical strategy of Beerli, Weiss, Zilibotti, and Zweimüller (2020), who study demand forces in technical change, employing an identification strategy to address the endogeneity of market size. Using cross-country datasets on consumption durables and productivity, we find robust evidence that market size for advanced goods has a significantly larger positive impact on productivity growth than market size for basic goods, supporting asymmetric externalities in our framework. In the quantitative analysis, we calibrate the model by matching key moments to their empirical counterparts. Even though our model features no friction—resources can move freely to their best use, our calibrated model shows that, purely due to the presence of dynamic production externalities, the two-sided policy intervention that corrects the externalities would yield sizable welfare gains and accelerate industrial upgrading.

Related Literature. There is a growing consensus in the empirical literature that the demand side, especially market size, is important for the pathway to successful development (see, e.g., Beerli et al. (2020), Goldberg and Reed (2023), Rodrik (2023), and Wen (2016)). In particular, market size determines the scale of learning-by-doing in production as well as the strength of incentives for innovation. Learning-by-doing is especially crucial for the cost-reduction type of technological progress (see, e.g., Barwick et al. (2025)), which is prevalent in developing countries.

Building on Stokey (1988) and Matsuyama (2002), our paper incorporates nonhomothetic preferences and learning-by-doing into a growth framework. Engel's law is linked to the Schmookler effect, whereby income elasticity differences across sectors affect relative market sizes, which in turn affect sectoral differences in productivity growth rates. A small body of prior literature studies the link, including Matsuyama (1992a, 2002), Zweimüller (2000), Foellmi and Zweimüller (2006), Bop-

part and Weiss (2013), and Oberfield (2023), focusing mainly on the relationship between inequality and growth (see also the recent related work by Bergman, Jaimovich, and Saporta-Eksten (2025)). Our focus is on dynamics with capital accumulation and optimal development policies.[5]

Our paper is related to the literature on structural change that emphasizes the role from the "demand side" through nonhomothetic preferences (e.g., Kongsamut, Rebelo, and Xie (2001), Matsuyama (2009), Buera and Kaboski (2012), Boppart (2014), Comin, Lashkari, and Mestieri (2021)).[6] Notably, Boppart (2014) and Comin, Lashkari, and Mestieri (2021) introduce non-vanishing nonhomothetic preferences into their structural estimations and find support for the presence of strong nonhomotheticity, with income effects being as important as substitution effects or even the main contributor to structural transformation. The hierarchical-preference utility used in our model is one type of nonhomothetic preferences. This literature typically treats technology as exogenous and independent of the demand side; in other words, there is Engel's law but no Schmookler effect.

The work of Foellmi and Zweimüller (2008), which studies structural change with infinite consumer goods introduced sequentially using hierarchical preferences, is closely related to our work. In their model extension, the authors also consider demand affecting the production side through the channel of new products introduced via endogenous R&D. Our paper is in a similar vein, but differs from their contribution in important respects. The source of growth in their paper is R&D à la Romer (1990) and Grossman and Helpman (1991) while our paper emphasizes externalities and spillovers (learning-by-doing) and focuses on the cost-reduction type of technological progress. Their model assumes identical technologies in different consumption-goods sectors, whereas studying endogenous productivity growth differences across sectors is a key focus of our paper. Our two-sector structure à la Rebelo (1991) also distinguishes between investment and consumption goods technologies. These differences manifest in unique policy implications of our model.

Our paper contributes to the literature that explicitly studies the role of government in accelerating development in the presence of market frictions or externalities. In the seminal work on "big push" of Murphy, Shleifer and Vishny (1989), the existence of multiple equilibria stemming from pecuniary externalities justifies policy intervention (see also Krugman (1991), Matsuyama (1992b, 1995)). By emphasizing supply-side constraints in terms of financial frictions (see the review by Banerjee and Duflo (2005)), Itskhoki and Moll (2019) develop a model of optimal development policies with financial frictions. Because of the presence of a dynamic pecuniary externality, the optimal intervention involves pro-business policies in the early stages and pro-worker policies in the

[5] Our model is also broadly related to the directed technical change literature, which emphasizes the role of market size in shaping the direction of innovation (e.g., Acemoglu (1998, 2002), Acemoglu and Zilibotti (2001)). The recent work of Comin, Lashkari, and Mestieri (2025) on structural transformation of innovation has no capital.

[6] Supply-side theories in the structural change literature emphasize the role of cross-sectional differences in productivity growth rates and capital intensities, but such differences are treated as exogenous (e.g., Baumol (1967), Ngai and Pissarides (2007), Acemoglu and Guerrieri (2008)). Relatedly, Ju, Lin, and Wang (2015) propose a theory of endowment-driven structural change.

long run. In our paper, with a frictionless economy, the policy is to address demand-side problems and the source of inefficiency is the dynamic production (technological) externalities, differing from and complementing their work. An important distinction from their paper is that ours has direct implications for industrial policy, suggesting that the government should target emerging industries rather than those that are too mature or too premature. The recent work of Aghion et al. (2024) studies industrial policy in a "big-push" production network with complementarities and multiple steady states, highlighting the need for a policy portfolio to implement the social optimum.

Our finding is broadly consistent with Acemoglu, Aghion, and Zilibotti's (2006) finding that economies far behind the world technology frontier should adopt an investment-led model and switch to an innovation-led growth model as they approach the frontier. The policy implication of enhancing capital accumulation in the early stages of development in our model is in line with the finding of government promotion of upstream heavy manufacturing sectors in the early stages in Liu's (2019) work on industrial policy. Jaimovich and Rebelo (2017) show non-linear effects of taxation on growth, aligning with our implication that right policies shape long-run growth. In terms of methodology, our paper is related to recent theoretical advances in decomposing externalities and assessing aggregate welfare in general equilibrium (Lorenzoni (2008), Dávila and Korinek (2018), Dávila and Schaab (2023)) and designing optimal policies (Dávila and Walther (2022)).

The paper is organized as follows. Section 2 presents the main model. Section 3 considers externalities and spillovers in the investment goods sector. Section 4 provides empirical evidence and quantitative analyses. Section 5 conducts model extensions. Section 6 concludes.

2 The Model

We first set up the model and then analyze it, focusing on consumption externalities.

2.1 Model Setting

Time is continuous, $t \in [0, +\infty)$. There are two types of goods: capital goods and consumption goods. Capital is used to produce both consumption goods and new capital. Specifically, we consider a two-sector AK economy à la Rebelo (1991): one sector uses capital to produce consumption goods while the other sector uses capital to produce investment goods. We choose the investment good as the *numeraire* in the economy. Factor and product markets are competitive. We use lowercase letters to denote individual variables and uppercase letters to denote aggregate variables.

Households. There is a continuum of households with unit mass. A representative household

with initial capital $k(0)$ at time $t=0$ maximizes its lifetime utility

$$\max \int_0^{\infty} \exp(-\rho t)\, u(c(t))\, dt, \tag{1}$$

where $u(c(t)) = (1-\gamma)^{-\hat{\sigma}} \frac{c(t)^{1-\hat{\sigma}}}{1-\hat{\sigma}}$ and $c(t)$ denotes the consumption bundle at time t. A consumption bundle consists of a continuum of varieties indexed by $j \in (0, +\infty)$,[7] namely,

$$c(t) = \int_0^{+\infty} j^{-\gamma} c(j,t)\, dj, \;\; c(j,t) \in \{0,1\}, \text{ with } \gamma \in (0,1), \qquad (\textbf{Assumption 1}) \tag{2}$$

where $c(j,t)$ is a "zero-one" choice for tractability and analytical convenience as in Matsuyama (2002), and $j^{-\gamma}$ represents a hierarchy factor. The non-homothetic preference in (2) follows Foellmi and Zweimüller (2006, 2008) and Bertola, Foellmi and Zweimüller (2014). That is, the consumption bundle $c(t)$ is an aggregation of utility on consumption varieties, weighted with a hierarchy factor. Let $\Omega(t) = \{j \in \mathbb{R}^+ | c(j,t) = 1\}$, the set of varieties that are actually consumed at time t.

The household's budget constraint is

$$\dot{a}(t) = r(t)\, a(t) - E(t), \;\; E(t) \equiv \int_0^{+\infty} p(j,t)\, c(j,t)\, dj, \tag{3}$$

where $a(t)$ denotes the household's asset (capital) holdings,[8] $r(t)$ is the market rate of return on assets, $E(t)$ is total expenditure, and $p(j,t)$ is the price of variety j.

Consumption Goods Producers. For each consumption variety j, there is a representative firm which rents capital from households and produces using a constant-returns-to-scale technology: $y(j,t) = \frac{1}{\varphi(j|J(t))} k(j,t)$, where $y(j,t)$ is the output of variety j and $\varphi(j|J(t))$ denotes the unit capital requirement (with $J(t)$ defined below). The firm chooses output to maximize profits:

$$\max_{y(j,t) \geq 0} \{p(j,t)\, y(j,t) - R(t)\, \varphi(j|J(t))\, y(j,t)\}, \tag{4}$$

where $R(t)$ is the rental rate of capital. The optimality condition implies $y(j,t)\,(p(j,t) - R(t)\,\varphi(j|J(t))) = 0$, so either $p(j,t) = R(t)\,\varphi(j|J(t))$ in which case $y(j,t)$ is indeterminate or $p(j,t) < R(t)\,\varphi(j|J(t))$ in which case $y(j,t) = 0$.

The unit cost function $\varphi(j|J(t))$ characterizes the production technology for consumption goods. For empirical relevance and analytical convenience, we assume:

$$\varphi(j|J(t)) = \begin{cases} 1-\lambda & \text{if } j \leq J(t) \\ 1 & \text{if } j > J(t) \end{cases}, \qquad (\textbf{Assumption 2'}) \tag{5}$$

[7] As in Matsuyama (2002), good 0 can be food, while goods $j > 0$ are manufacturing goods.

[8] Implicitly, there is an asset market in which households deposit their capital with a competitive representative financial intermediary which rents out capital to firms, and households own the financial intermediary.

where $J(t)$ represents the "*(domestic) aggregate technology frontier*" at time t. That is, there are only two levels of productivity. A variety $j > J(t)$ can be produced with the domestic *immature* technology at the full cost 1, whereas a variety $j \leq J(t)$ can be produced with the domestic *mature* technology at the reduced unit cost $1-\lambda$.

For the purposes of general analysis and quantification later on, we present here a generalized setting with a "smooth" function of $\varphi(j|J(t))$:

$$\varphi(j|J(t)) = 1 - \lambda\left[1 + \kappa\varepsilon\left(\frac{J(t)}{j}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{1}{\varepsilon-1}} \quad \text{with } \varepsilon \in (0,1). \qquad (\textbf{Assumption 2}) \qquad (6)$$

Note that the limiting case as $\varepsilon \to 0^+$ of (6) converges to (5).[9] The specific function form of $\varphi(j|J(t))$ in (6) is to obtain a CES-like cost function (see (15) later). In (6), $J(t)$ can be interpreted as an *index* of aggregate technology in the consumption goods sector at time t or the sector's cumulative learning experience ("*experience capital*"), as in Stokey (1988) and Young (1993). The full unit cost of production for any variety is 1; however, due to the aggregate technological progress $J(t)$, the unit cost is reduced by $\lambda\left[1 + \kappa\varepsilon\left(\frac{J(t)}{j}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{1}{\varepsilon-1}}$ for variety j. The term $\left[1 + \kappa\varepsilon\left(\frac{J(t)}{j}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{1}{\varepsilon-1}}$ is decreasing in j, increasing in $J(t)$, and bounded within $[0,1]$. The parameter ε, which will be shown to correspond to the elasticity of substitution in a CES function, can be interpreted as the speed and concentration of learning-by-doing spillovers, i.e., how quickly and intensively industries around $j = J(t)$ will experience cost reductions as $J(t)$ advances.[10] The parameter $\kappa > 0$ is a scaling term. Figure 1 plots $\varphi(j|J(t))$.

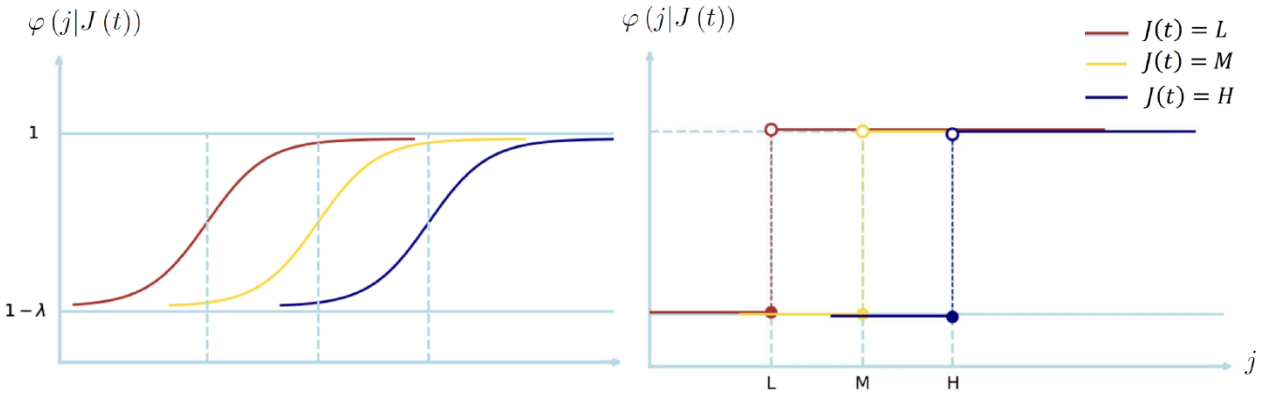


Figure 1: The unit cost function $\varphi(j|J(t))$ (Left: ε bounded away from zero; Right: $\varepsilon \to 0^+$)

[9] Note that the simpler form $\varphi(j|J(t)) = 1 - \lambda\left[1 + \left(\frac{J(t)}{j}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{1}{\varepsilon-1}}$ would yield the limiting case as $\lim_{\varepsilon\to 0^+} \varphi(j|J(t)) = 1 - \frac{1}{2}\lambda$ for $j = J(t)$, which is less relevant and would complicate the analysis in Section 2.4.

[10] When $\varepsilon < 1/2$, the function $\varphi(j|J)$ is "S"-shaped as shown in Figure 1, which is the case our quantitative analysis will focus on in Section 4.

Hierarchical Demand. Given the preferences and the production function, it is easy to show that hierarchical demand for consumption varieties follows. Formally, at any time t, if some $j \in \Omega(t)$, then $j' \in \Omega(t)$ for all $0 < j' < j$. We can therefore write $\Omega(t) = (0, n(t)]$, where $n(t)$ denotes the cut-off variety (or "consumption frontier") that satisfies $c(n^-(t), t) = 1$ and $c(n^+(t), t) = 0$. Let $N(t)$ denote the average $n(t)$ across households in the economy.

Technology Process. In the same spirit of Stokey (1988) and Matsuyama (2002) and many other studies, we model *demand-induced* technological progress by assuming

$$\dot{J}(t) = \mathcal{J}(N(t), J(t)) = \begin{cases} bJ(t) + \phi_0 N(t)^\alpha & \text{if } N(t) \leq J(t) \\ bJ(t) + \phi N(t)^\alpha & \text{if } N(t) > J(t) \end{cases}, \text{ with } \alpha \in (0,1) \quad \textbf{(Assumption 3')} \tag{7}$$

where parameters satisfy $\phi > \phi_0 \geq 0$. Here, the parameter b represents the natural or exogenous growth rate of technology, while the $N(t)$-term captures the endogenous growth component, reflecting demand-induced technological progress.

Assumption 3' states that $N(t)$ positively affects $\dot{J}(t)$ and that this effect is stronger when $N(t)$ is above than below $J(t)$. The asymmetry ($\phi > \phi_0$) is intended to capture the idea that $N(t)$ has a stronger "pull" effect on $J(t)$ when $N(t)$ is above $J(t)$.[11] This feature aligns with studies emphasizing that the production of basic, subsistence goods (low indexed varieties in our model) generates minimal production externalities, whereas the production of more advanced consumption goods generates significant externalities and spillovers (see, e.g., Matsuyama (1992a) and Lucking et al. (2019)). We will show that the asymmetric learning assumption ($\phi > \phi_0$) is quantitatively important for our model's conclusion, but the model conclusion still holds qualitatively under symmetric learning ($\phi = \phi_0 > 0$).

The restriction $\alpha < 1$ in Assumption 3' implies diminishing returns to learning-by-doing, consistent with the literature (see, e.g., Young (1991, 1993), Acemoglu (2009)). In our model, this restriction ensures that $J(t)$ does not grow faster than $N(t)$ in the long run and thus does not eventually surpass $N(t)$. Conceptually, $\alpha < 1$ captures the idea that learning-by-doing as a force behind growth has limits and boundaries and will exhaust its strength when the economy gets sufficiently close to the world technology frontier (see, e.g., Young (1991) and Zilibotti (2017)).

We will microfound the law of motion of (7) in Section 5.1. The core mechanism is that demand can induce innovation and thus pull the (domestic) aggregate technology forward, due to the scale of learning-by-doing (Lucas (1988)) and/or the strength of incentives to innovate (Schmookler (1966)).[12] The assumption of learning-by-doing is the simplest modeling device, as noted by many

[11] Inter-industry spillover of learning by doing has been incorporated into models by Stokey (1988), Young (1991), and many others. Cai and Li (2019), among others, document evidence of inter-sectoral knowledge spillovers.

[12] Innovation in our model is more about the type that adopts foreign technology for domestic use (Krugman (1979), Acemoglu and Zilibotti (2001)) or reduces the production costs of existing varieties (Aghion and Howitt (1992), Grossman and Helpman (1991)), and less about the type that creates new varieties (Romer (1990)).

authors (e.g., Acemoglu (2009)). What is crucial in our context is the asymmetry in knowledge generation between goods $j < J(t)$ and goods $j > J(t)$: producing the latter generates more knowledge. This can arise because production with immature technology naturally generates more knowledge as a by-product (see, e.g., the evidence in Lamoreaux, Raff, and Temin (1998)).

For the purposes of general analysis and quantification later on, we present here a generalized setting with a "smooth" function of $\mathcal{J}(N(t), J(t))$:

$$\dot{J}(t) = \mathcal{J}(N(t), J(t)) = bJ(t) + \beta(N(t), J(t)) \cdot N(t)^{\alpha}, \text{ with } \alpha \in (0,1) \quad (\textbf{Assumption 3}) \tag{8}$$

where $\beta(N, J)$ is a (modified) logistic function

$$\beta(N, J) = \phi_0 + \frac{\phi - \phi_0}{1 + \frac{\nu}{\zeta}\exp\left(-\frac{N-J}{\zeta}\right)}, \text{ with } \zeta > 0, \tag{9}$$

and parameters satisfy $\nu > 0$ and $\phi > \phi_0 \geq 0$. The logistic form $\beta(N, J)$ follows the technology diffusion literature (e.g., Griliches (1957), Benhabib et al. (2014), Perla and Tonetti (2014)). The limiting case as $\zeta \to 0^+$ of $\beta(N, J)$ converges to a step function, yielding (7).[13]

Investment Goods Producers. There is a representative firm which rents capital from households to produce investment goods using a constant-returns-to-scale technology: $i(t) = zk_I(t)$, where $i(t)$ denotes investment goods output and z is the productivity (TFP). The firm chooses output to maximize profits:

$$\max_{i(t)\geq 0}\{p_I(t)\, i(t) - R(t)\,(i(t)/z)\}. \tag{10}$$

Note that we choose the investment good as the numeraire in the economy, so that $p_I(t) \equiv 1$ for all t. The optimality condition implies that for all t

$$R(t) = z. \tag{11}$$

Capital Accumulation. Let $I(t)$ denote the aggregate amount of investment goods produced at time t. Aggregate capital accumulates according to

$$\dot{K}(t) = I(t) - \delta K(t), \tag{12}$$

where δ is the capital depreciation rate after productions.

We also assume that

$$k(0) \text{ is sufficiently low relative to } J(0). \qquad (\textbf{Assumption 4})$$

[13] The simpler form $\beta(N, J) = \phi_0 + \frac{\phi - \phi_0}{1+\exp\left(-\frac{N-J}{\zeta}\right)}$ does not yield (7) in the limiting case $\zeta \to 0^+$.

Further specification of this assumption will be given in Proposition 1. Reasonably, for a country even with a very low level of initial capital stock $k(0)$, its technology stock $J(0)$ can be relatively better because of the existence of international technology (know-how) spillovers (e.g., Krugman (1979)). (A post-war economy may also be the case.)

2.2 Characterization of Equilibrium

We characterize the equilibrium under the general assumptions 2 and 3.

Within-period Aggregation. Since $\Omega(t) = (0, n(t)]$, and $c(j,t) = 1$ for $j \leq n(t)$ and 0 for $j > n(t)$, the consumption bundle $c(t)$ is calculated as

$$c(t) = \int_0^{n(t)} j^{-\gamma} c(j,t)\, dj = \frac{1}{1-\gamma} n(t)^{1-\gamma}. \tag{13}$$

The objective function in (1) can hence be rewritten as $\max \int_0^\infty \exp(-\rho t) \left(\frac{1}{1-\sigma} n(t)^{1-\sigma}\right) dt$, where $\sigma \equiv 1 - (1-\gamma)(1-\hat{\sigma})$. Moreover, the expenditure function at time t is calculated as

$$E(t) = \int_0^{n(t)} p(j,t)\, c(j,t)\, dj = R(t) \int_0^{n(t)} \varphi(j|J(t))\, dj \equiv R(t)\, \mathbf{e}(n(t); J(t)), \tag{14}$$

where

$$\mathbf{e}(n(t); J(t)) = \int_0^{n(t)} \varphi(j|J(t))\, dj = n(t) - \lambda \left[n(t)^{\frac{\varepsilon-1}{\varepsilon}} + (\kappa\varepsilon) J(t)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}. \tag{15}$$

Note that $\mathbf{e}$ is the cost function denominated in units of physical capital, whereas E is the expenditure function denominated in the numeraire (the investment good). The expression in (15) is intuitive. Under the full unit cost 1 in (6), total cost would be $n(t)$; however, technological progress $J(t)$ reduces this cost by $\int_0^{n(t)} \lambda \left[1 + (\kappa\varepsilon)\left(\frac{J(t)}{j}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{1}{\varepsilon-1}} dj = \lambda \left[n(t)^{\frac{\varepsilon-1}{\varepsilon}} + (\kappa\varepsilon) J(t)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}$.

Notice that the term $\mathcal{D}(n(t), J(t)) \equiv \left[n(t)^{\frac{\varepsilon-1}{\varepsilon}} + (\kappa\varepsilon) J(t)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}$ is a CES aggregator, where $n(t)$ and $J(t)$ exhibit gross complementarity under $\varepsilon \in (0,1)$. In fact, $\mathcal{D}_J(n,J) = (\kappa\varepsilon)\left[(n/J)^{\frac{\varepsilon-1}{\varepsilon}} + (\kappa\varepsilon)\right]^{\frac{1}{\varepsilon-1}}$ is monotonically increasing in n/J. In the limiting case $\varepsilon \to 0^+$, $\mathcal{D}(n(t), J(t)) = \min(n(t), J(t))$ (a Leontief function),[14] and $\mathcal{D}_J(n,J)$ equals 0 when $n/J < 1$ and 1 when $n/J > 1$. These properties of $\mathcal{D}(n,J)$ (and hence $\mathbf{e}(n;J)$) will be useful later.

Market Clearing. Asset market clearing implies that for all t

$$a(t) = k(t). \tag{16}$$

[14] Although ε enters into the weight of CES, this can be verified since $\lim_{\varepsilon \to 0^+} (\kappa\varepsilon)^{\frac{\varepsilon}{\varepsilon-1}} = 1$ by L'Hôpital's rule.

Market clearing for consumption varieties implies that $c(j,t) = y(j,t)$ for all j and t. Market clearing of capital goods implies that for all t

$$K_I(t) = I(t)/z, \;\; K_C(t) = \mathbf{e}(n(t); J(t))$$

and

$$K_C(t) + K_I(t) = K(t), \tag{17}$$

where $K_I(t)$ and $K_C(t)$ denote the aggregate supply of capital to the investment and consumption goods sectors, respectively.

Definition of Equilibrium. We define the competitive equilibrium (CE) and the social planner problem (SP).

Definition 1 *The CE consists of individual choices* $\mathcal{I} = \{a(t), c(j,t), y(j,t), i(t)\}_{j\in(0,n(t)],t=0}^{\infty}$, *aggregate allocations* $\mathcal{A} = \{K(t), C(j,t), Y(j,t), I(t)\}_{j\in(0,n(t)],t=0}^{\infty}$, *and prices* $\mathcal{P} = \{p(j,t), R(t), r(t)\}_{j\in(0,n(t)],t=0}^{\infty}$ *such that for some given* k_0 *and* J_0 *the following conditions are satisfied:*

1) *Given* $\mathcal{A}$ *and* $\mathcal{P}$, $\mathcal{I}$ *solves households' problem (1), consumption goods producers' problem (4), and investment goods producers' problem (10);*

2) *The aggregate allocations are consistent with the individual choices:* $\mathcal{A} = \mathcal{I}$;

3) *Aggregate capital evolves according to (12) and* $J(t)$ *is determined by the dynamics (8);*

4) *At each* t, $p(j,t)$ *clears the* j*-th variety market such that* $c(j,t) = y(j,t)$, $R(t)$ *clears the capital market such that (17) is satisfied, and* $r(t)$ *clears the asset market such that (16) is satisfied.*

In contrast, the SP is to characterize the allocation of resources that maximizes the utility of the normative representative household.

Based on *(4)*, *(10)*, and *(12)*, it is easy to find that in CE, for all t, $p(j,t) = R(t)\varphi(j|J(t))$ for $j \in (0, n(t)]$, $R(t) = z$, and

$$r(t) = R(t) - \delta. \tag{18}$$

Plugging *(14)*, *(16)*, *(11)*, *and (18) into (3) yields* $\dot{k}(t) = (z-\delta)k(t) - z\mathbf{e}(n(t); J(t))$. The equilibrium can hence be stated in an alternative way.

Definition 1 ***(alternative)*** *The competitive equilibrium is a fixed-point problem between the two*

processes $\{J(t)\}_{t=0}^{\infty}$ *and* $\{N(t)\}_{t=0}^{\infty}$*, given by*

$$\max_{\{n(t)\}_{t=0}^{\infty}} \int_0^{\infty} \exp(-\rho t)\left(\frac{1}{1-\sigma} n(t)^{1-\sigma}\right) dt \tag{19}$$
$$\text{subject to} \quad \dot{k}(t) = (z-\delta)\,k(t) - z\mathbf{e}(n(t); J(t)) \text{ and some given } k(0) > 0$$

and

$$\{J(t)\}_{t=0}^{\infty}: \ \dot{J}(t) = \mathcal{J}(N(t), J(t)) \text{ and some given } J(0) > 0. \tag{20}$$

The social planner problem is with the same optimization program (19), but taking equation (20) as a control state equation rather than as exogenous.

In CE, standing at $t = 0$, individual households take the entire process $\{J(t)\}_{t=0}^{\infty}$ as exogenously given and choose their consumption path $\{n(t)\}_{t=0}^{\infty}$, while the process of $J(t)$ in turn depends on the process of aggregate $N(t)$.

2.3 General Analysis on the Model

We analyze the model under the general assumptions 2 and 3 with smooth functions of $\varphi(j|J)$ and $\mathcal{J}(N, J)$. Under the smooth functions, the Euler equation is binding everywhere in the equilibrium path, so a general and abstract analysis can be conducted, outlining the essential mechanism.

Characterizing Dynamic Externality. The core mechanism of our model concerns the dynamic externality induced by consumption. For the CE optimization problem in Definition 1 (alternative), the current-value Hamiltonian is formed as

$$\mathcal{H}^{CE} = \frac{1}{1-\sigma} n^{1-\sigma} + m\left[(z-\delta)\,k - z\mathbf{e}(n; J)\right],$$

where m is the costate variable, corresponding to $m = V_k$ (i.e., the marginal valuation of capital on the value function V). The optimality conditions imply $n^{-\sigma} - mz\mathbf{e}_n(n; J) = 0$ and $\frac{\dot{m}}{m} = -(z-\delta-\rho)$, so we obtain the consumption Euler equation for the CE:

$$\left.\frac{\dot{n}}{n}\right|_{\text{CE}} = \frac{1}{\sigma}(r(t) - \rho) - \frac{1}{\sigma}\frac{d\log \mathbf{e}_n(n; J)}{dt}, \tag{21}$$

where $r(t) = z - \delta$ is the rate of return on assets (recalling (11) and (18)). Intuitively, $\dot{n}/n$ depends on the intertemporal interest rate as well as the growth rate of the marginal cost of consumption. Similarly, for the SP optimization problem, the current-value Hamiltonian is formed as

$$\mathcal{H}^{SP} = \frac{1}{1-\sigma} n^{1-\sigma} + m\left[(z-\delta)\,k - z\mathbf{e}(n; J)\right] + \mu\mathcal{J}(n, J),$$

where μ is the costate variable, corresponding to $\mu = V_J(K, J)$. The optimality conditions imply

$n^{-\sigma} - mz\mathbf{e}_n(n;J) + \mu\mathcal{J}_n(n,J) = 0$, $\frac{\dot{m}}{m} = -(z-\delta-\rho)$, and $\dot{\mu} = z\mathbf{e}_J(n;J)\,m + (\rho - \mathcal{J}_J(n,J))\,\mu$. Rewriting $n^{-\sigma}\left(1+\mu\mathcal{J}_n(n,J)/n^{-\sigma}\right) - mz\mathbf{e}_n(n;J) = 0$ and substituting for m yields the consumption Euler equation for the SP:

$$\left.\frac{\dot{n}}{n}\right|_{\text{SP}} = \frac{z-\delta-\rho}{\sigma} - \frac{1}{\sigma}\frac{d\log \mathbf{e}_n(n;J)}{dt} + \frac{1}{\sigma}\underbrace{\frac{d\log\left(1+\overbrace{\mathcal{J}_n(n,J)\,\mu/n^{-\sigma}}^{\text{externality of additional consumtion at } t}\right)}{dt}}_{\textbf{growth of externality}}. \tag{22}$$

Comparing (22) with (21), we see the key difference in the consumption growth rate ($\frac{\dot{n}}{n}$) between CE and SP, which arises due to the existence of externalities. Specifically, individual households do not internalize the effects of $n(t)$ on $\dot{J}(t)$ (equation (20)). Importantly, it is the *dynamics* (growth) of the externality (characterized by the term $d\log\left(1+\mathcal{J}_n(n,J)\,\mu/n^{-\sigma}\right)/dt$) that matters. Note that the externality of additional consumption at time t has two components: i) the term $\mathcal{J}_n(n(t),J(t))$, i.e., how much $\dot{J}(t)$ can be increased by increasing $n(t)$, and ii) the term $\mu(t)/n(t)^{-\sigma}$, i.e., how much *equivalent* consumption at t can be increased by increasing $J(t)$.[15]

Optimal Policy Intervention. Consider that the government provides a subsidy to the savings rate of return in CE at rate $\mathcal{S}(t)$ (that is, households receive the asset return rate $r(t)(1+\mathcal{S}(t))$). Under this scheme of intervention, the new Euler equation for the CE becomes $\left.\frac{\dot{n}}{n}\right|_{\text{CE with tax}} = \frac{r(t)(1+\mathcal{S}(t))-\rho}{\sigma} - \frac{1}{\sigma}\frac{d\log \mathbf{e}_n(n(t);J(t))}{dt}$, where $r(t) = z-\delta$ still holds true. The implementation of the SP allocation, $\left.\frac{\dot{n}}{n}\right|_{\text{CE with tax}} = \left.\frac{\dot{n}}{n}\right|_{\text{SP}}$, then implies

$$\mathcal{S}(t) = \frac{1}{z-\delta}\underbrace{\frac{d\log\left(1+\mathcal{J}_n(n,J)\,\mu/n^{-\sigma}\right)}{dt}}_{\textbf{growth of externality}}. \tag{23}$$

The result of (23) is crucial. It clearly shows that the subsidy rate $\mathcal{S}(t)$ is positive (negative) if and only if the externality increases (decreases) over time. In other words, if the government suppresses consumption ($\mathcal{S}(t) > 0$) first and simulates consumption ($\mathcal{S}(t) < 0$) later, it must mean that the externality first increases and then decreases over time. Figure 2 plots the dynamics of some key variables using the calibrated parameters in the quantitative exercise in Section 4.2.

Explaining Mechanism. We now explain why the externality first increases and then decreases over time in our model. Basically, we show that both the term $J_n(n,J)$ and the term μ in (23) exhibit a hump-shaped trajectory, as illustrated in Figure 2.

First, consider the pattern of $J_n(n,J)$. The *asymmetry learning* (i.e., $\phi_0 < \phi$) in Assumption 3 implies that $J_n(n,J)$ is smaller when $n < J$ than when $n > J$. Dynamically, this means that

[15] In the term $\mu(t)/n(t)^{-\sigma}$, the denominator $n(t)^{-\sigma}$ is the marginal utility of consumption at time t. So the term corresponds to the conversion of an increase in lifetime utility into consumption units at time t.

$J_n(n, J)$ increases over time in the initial phase. Moreover, the condition $\alpha < 1$ in Assumption 3 implies that $\lim_{t\to\infty} J_n(n(t), J(t)) = 0$. These together imply that $J_n(n, J)$ follows a hump-shaped pattern over its lifetime.

Second, we examine the dynamics of $\mu \equiv V_J(K, J)$. The law of motion for μ is given by

$$\dot{\mu} = z\mathbf{e}_J(n; J)\, m + (\rho - \mathcal{J}_J(n, J))\, \mu,$$

where $m \equiv V_k$ with $\frac{\dot{m}}{m} = -(z - \delta - \rho)$. The key term $\mathbf{e}_J(n; J) = -\lambda(\kappa\varepsilon)\left[(n/J)^{\frac{\varepsilon-1}{\varepsilon}} + (\kappa\varepsilon)\right]^{\frac{1}{\varepsilon-1}}$, representing the degree of cost reduction due to an increase in J, has the property that $-\mathbf{e}_J(n, J)$ is increasing in n/J.[16] Dynamically, this means that $-\mathbf{e}_J(n; J)$ increases over time. However, $m(t)$ diminishes over time, with $\lim_{t\to\infty} m(t) = 0$. Consequently, μ also follows a hump-shaped pattern. The economic intuition behind the dynamics of μ is *capital-technology complementarity*. When $n < J$, this means that capital is relatively scarce and insufficient to fully utilize the available technology—technology is 'redundant'. In this phase, the marginal value of technology is low (i.e., $-\mathbf{e}_J(n; J)$ is low), which in turn implies that μ is also low and increasing over time. When $n > J$, technology is no longer redundant and can be fully utilized, raising $-\mathbf{e}_J(n; J)$ and thus μ. However, since m will eventually diminish to zero, $z\mathbf{e}_J(n; J)\, m$ will approach zero, and so will μ.

The joint hump-shaped evolution of the two components gives rise to the hump-shaped pattern of externality and its two-sided growth rate.[17] Since there are two joint forces, our result on two-sided intervention can hold qualitatively even if the first force of asymmetric learning is completely shut down (i.e., $\phi_0 = \phi > 0$ in Assumption 3); see the right-hand panel of Figure 2.

[16] This property is sufficient, but not necessary, for the model conclusion. A weaker condition that $-\mathbf{e}_J(n, J)$ is higher when $n > J$ than when $n < J$ is sufficient.

[17] We will show that $\mu(t)/n(t)^{-\sigma}$ converges to a constant as $t \to \infty$. Consequently, the externality term $\mathcal{J}_n(n, J)\, \mu/n^{-\sigma}$ approaches 0 as $t \to \infty$.

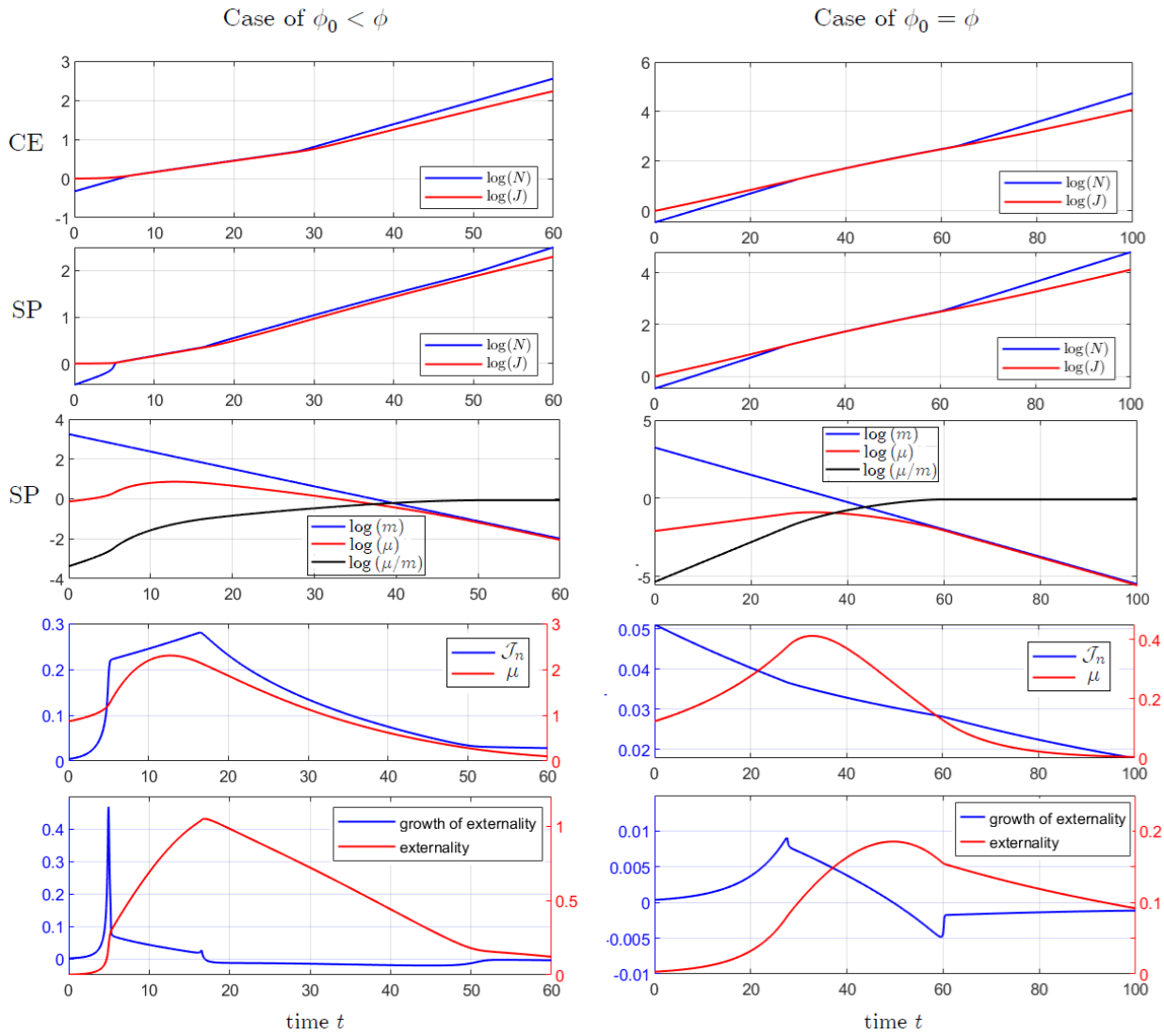


Figure 2: Understanding the model's mechanism

Note: This figure plots the equilibrium paths of some key variables. The last row shows the level and the growth rate of the externality, where the former is defined as $J_n(n,J)\,\mu/n^{-\sigma}$ and the latter as $d\log\left(1+\mathcal{J}_n(n,J)\,\mu/n^{-\sigma}\right)/dt$. The optimal subsidy $\mathcal{S}(t)$ is the growth rate of the externality, multiplied by the constant $1/(z-\delta)$. The left-hand panel shows the results under the calibrated parameter values listed in Table 4 in Section 4.2 and the right-hand panel shows the results under the same set of parameter values except that ϕ_0 is set equal to ϕ.

Illustration of Intuition. Figure 3 illustrates the intuition behind the increase in externality in the early development phase. In the 1980s, a country had a low capital stock and its consumption frontier was limited to basic goods like bikes, but its technology stock was relatively high, as reflected by the domestic technology frontier being cars. By 2023, the consumption frontier had advanced to home robots, and the technology frontier had progressed to cell phones. Reasonably,

the production externalities and spillovers from bikes (low-end goods) to cars (high-end goods) were modest, whereas those from home robots to cell phones were substantial in 2023, i.e., "asymmetric spillovers". Thus, from the standpoint of the 1980s, the future externality was higher than the current externality (*the first force*—$\mathcal{J}_n$ increases over time). Moreover, even if consumption-induced externalities improved car production technology in the 1980s, such an improvement could not be utilized because the capital stock at that moment was too low, only enough to produce varieties up to bike. Technological advances had little value without enough capital to utilize them (*the second force*—$\mu = V_J(K, J)$ increases over time). Overall, from the 1980s perspective, the forward-looking growth of the externality is positive in (23). If individual households are indifferent in their intertemporal tradeoff between consuming a bike in the 1980s and a home robot in 2023 in CE, the planner can enhance welfare by encouraging households to consume less and accumulate capital in the 1980s. This enables greater consumption in 2023 and thus exploits the larger production externalities present then.

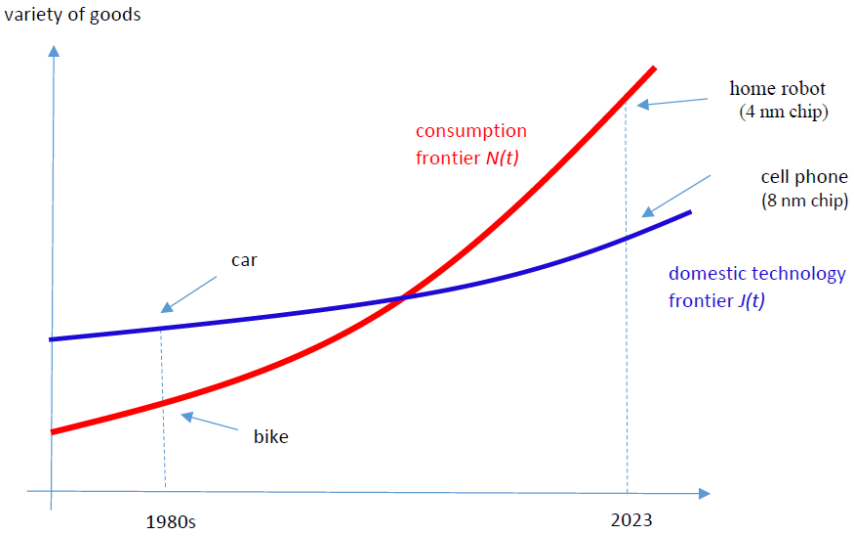


Figure 3: An illustration of dynamic externalities in the early development phase

2.4 Solving Equilibrium under Limiting Cases of Assumptions

We now solve the model in closed form under limiting cases of the assumptions, enabling us to characterize in closed form the full dynamics and optimal policy intervention.

Limiting Cases of Assumptions. First, for Assumption 2, we focus on the limiting case $\varepsilon \to 0^+$, which is Assumption 2'. We restate it here:

$$\varphi(j|J(t)) = \begin{cases} 1-\lambda & \text{if } j \leq J(t) \\ 1 & \text{if } j > J(t) \end{cases}. \qquad \textbf{(Assumption 2')} \tag{24}$$

So the cost function becomes

$$\mathbf{e}(n(t); J(t)) = n(t) - \lambda \min(n(t), J(t)) = \begin{cases} (1-\lambda) n(t) & \text{if } n(t) \leq J(t) \\ n(t) - \lambda J(t) & \text{if } n(t) > J(t) \end{cases}, \tag{25}$$

where function $\mathbf{e}$ is a piecewise linear function with a *kink* at $n(t) = J(t)$.

Second, for Assumption 3, we focus on the limiting case $\zeta \to 0^+$ with $\phi_0 = 0$, which is Assumption 3' with $\phi_0 = 0$, that is,

$$\dot{J}(t) = \mathcal{J}(N(t), J(t)) = \begin{cases} bJ(t) & \text{if } N(t) \leq J(t) \\ bJ(t) + \phi N(t)^{\alpha} & \text{if } N(t) > J(t) \end{cases}. \quad \textbf{(Assumption 3")} \tag{26}$$

Assumption 3" states that $N(t)$ has a positive impact on $\dot{J}(t)$ when $N(t) > J(t)$ and has no impact when $N(t) \leq J(t)$. This has an intuitive interpretation: the demand for and consequently the production of lower-end goods has no learning-by-doing spillovers on the production of higher-end goods, while significant spillovers exist in the opposite direction.

2.4.1 Solving the Competitive Equilibrium

Solving the competitive equilibrium is challenging because of the presence of *regime changes (kinks)* in the cost function $\mathbf{e}$ as well as in the technology process $J(t)$, coupled with the fixed-point problem. Nevertheless, we are able to solve the competitive equilibrium in closed form.

Proposition 1 ***(Competitive equilibrium)*** *Given the initial state* $(K(0) = K_0, J(0) = J_0)$*, when* K_0 *is low enough relative to* J_0 *such that* $K_0 < K^*(J_0)$*,*[18] *the equilibrium processes* $\{N(t), K(t), J(t)\}_{t=0}^{\infty}$ *are unique and have three regimes (stages) in* $t \in [0, \tau_1] \cup (\tau_1, \tau_2) \cup [\tau_2, +\infty)$*:*

- *for* $t \in [0, \tau_1]$
  $n(t) = n(0) \exp\left[\frac{(z-\delta)-\rho}{\sigma} t\right]$
  $J(t) = J(0) \exp(bt)$
  $\dot{k}(t) = (z-\delta) k(t) - z(1-\lambda) n(t)$ *with* $k(0)$ *given*
- *for* $t \in (\tau_1, \tau_2)$
  $n(t) = J(t)$
  $J(t) = J(0) \exp(bt)$
  $\dot{k}(t) = (z-\delta) k(t) - z(1-\lambda) J(t)$ *with* $k(\tau_1^+) = k(\tau_1^-)$
- *for* $t \in [\tau_2, +\infty)$

[18] If $K_0 \geq K^{**}(J_0)$ (defined in the proof), only the third stage exists. If $K^{**}(J_0) > K_0 \geq K^*(J_0)$, only the second and third stages exist. See the Appendix for details.

$$
\begin{aligned}
&n(t) = n(\tau_2)\exp\left[\tfrac{(z-\delta)-\rho}{\sigma}(t-\tau_2)\right] \;\; \textit{with } n\left(\tau_2^+\right) = n\left(\tau_2^-\right) \\
&\dot{J}(t) = bJ(t) + \phi N(t)^{\alpha} \;\; \textit{with } J\left(\tau_2^+\right) = J\left(\tau_2^-\right) \\
&\dot{k}(t) = (z-\delta)k(t) - z\left[n(t) - \lambda J(t)\right] \;\; \textit{with } k\left(\tau_2^+\right) = k\left(\tau_2^-\right)
\end{aligned}
,
$$

where $K^*(\cdot)$, $n(0)$, τ_1, *and* τ_2 *are endogenous and analytically solved, that is, 1) function* $K^*(\cdot)$ *is given in (A.8) in the appendix; 2) the two unknowns* $n(0)$ *and* τ_1 *are the solution of the joint equations:*

$$
\begin{cases}
k(\tau_1) = K^*\left(J(\tau_1)\right) \\
n(0)\exp\left(\frac{z-\delta-\rho}{\sigma}\tau_1\right) = J(\tau_1),
\end{cases}
$$

where $k(\tau_1) = \left(k(0) - n(0)\frac{z(1-\lambda)}{(z-\delta)-\frac{z-\delta-\rho}{\sigma}}\left[1-\exp\left[-\left((z-\delta)-\frac{z-\delta-\rho}{\sigma}\right)\tau_1\right]\right]\right)\exp\left[(z-\delta)\tau_1\right]$ *and* $J(\tau_1) = J(0)\exp(b\tau_1)$*; and 3)* $\tau_2 = \tau_1 + \frac{-\log(1-\lambda)}{(z-\delta-\rho)-\sigma b}$.

Proof: *See Appendix.*

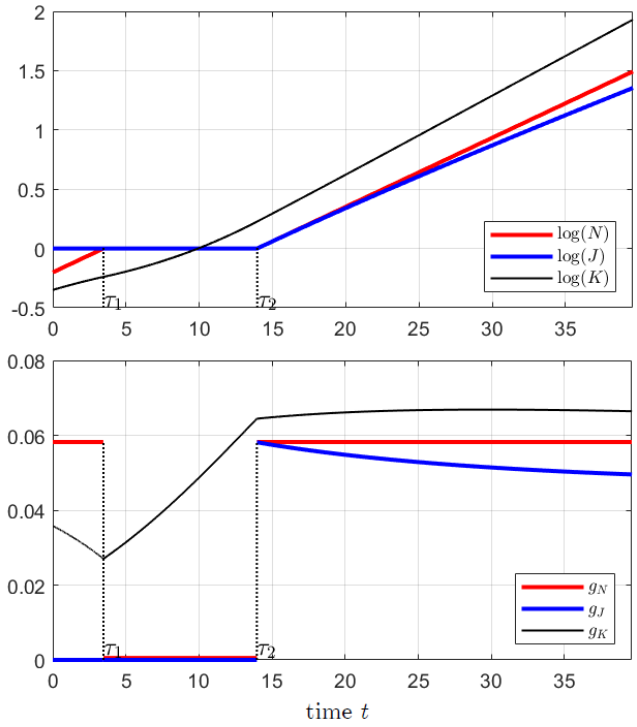


Figure 4: The competitive equilibrium processes $N(t)$, $K(t)$, and $J(t)$

Figure 4 illustrates the equilibrium processes $\{N(t), K(t), J(t)\}_{t=0}^{\infty}$ under the same set of parameter values as in Figure 2 (see Table 4 in Section 4.2) except for $\varepsilon \to 0^+$ and $\zeta \to 0^+$. A key feature of the processes is that the second stage is a "corner solution". Specifically, $N(t)$ initially grows at a constant rate $\frac{z-\delta-\rho}{\sigma}$, and then sticks to the technology frontier path $J(t)$ with the growth rate b, and then takes off and grows at a constant rate $\frac{z-\delta-\rho}{\sigma}$. That is, in the middle stage

between τ_1 and τ_2, the consumption $N(t)$ stays at the level of the technology frontier $J(t)$ for a while; until the capital accumulates and reaches to a sufficiently high level, the consumption $N(t)$ grows faster again. The "flat" part in the middle stage reflects that the kink at $n = J$ in the cost function $\mathbf{e}(n; J)$ causes a "corner solution" $(n = J)$ in the optimization over the Hamiltonian (see the proof); note that in Figure 4, the parameter value b has been set to 0. Importantly, the two thresholds τ_1 and τ_2 are endogenous and are functions of the initial state $(K(0), J(0))$.

Denote $g_N(t) \equiv \frac{\dot{N}(t)}{N(t)}$, $g_J(t) \equiv \frac{\dot{J}(t)}{J(t)}$, and $g_K(t) \equiv \frac{\dot{K}(t)}{K(t)}$. From Proposition 1, it is easy to obtain the following corollary.

Corollary 1 *The economy asymptotically approaches the steady state with* $\lim_{t\to\infty} g_N(t) = \lim_{t\to\infty} g_K(t) = \frac{(z-\delta)-\rho}{\sigma}$, $\lim_{t\to\infty} g_J(t) = \alpha\frac{(z-\delta)-\rho}{\sigma}$, $\lim_{t\to\infty} \frac{N(t)}{K(t)} = \frac{(z-\delta)-\frac{(z-\delta)-\rho}{\sigma}}{z}$, *and* $\lim_{t\to\infty} \frac{J(t)}{N(t)^\alpha} = \frac{\phi}{\alpha\frac{(z-\delta)-\rho}{\sigma}-b}$.

To ensure positive growth and household utility being bounded away from infinity, we impose the parameter conditions such that $\alpha\frac{(z-\delta)-\rho}{\sigma} > b$ and $\frac{(z-\delta)-\rho}{\sigma}(1-\sigma) < \rho$.

For simplicity and without loss of generality, we set $b = 0$ in the rest of this section.

2.4.2 Social Planner's Solution and Implementation

We solve the planner problem, compare the planner's allocation with the CE allocation, and analyze how the planner's allocation can be implemented in the decentralized economy. We show that a two-phase policy intervention with *opposite signs* at different phases is necessary to eliminate the inefficiency in the decentralized economy due to the presence of the dynamic externality.

Comparison between CE and SP. We provide details of the solution to the planner problem in the Appendix. Figure 5 plots the social planner's path compared to the CE path, with the same initial state (K_0, J_0) and parameter values as in Figure 4.

In Figure 5, the planner's optimal consumption path also has three stages (regimes).[19] It is initially below the CE's consumption path, but later above the CE's path, and eventually has the same steady-state growth rate as the CE given in Corollary 1. An important feature is that the duration of the second stage $(t \in [\tau_1, \tau_2])$ is shorter for SP than for CE.

[19] See the proof of Proposition 2 in the Appendix for details.

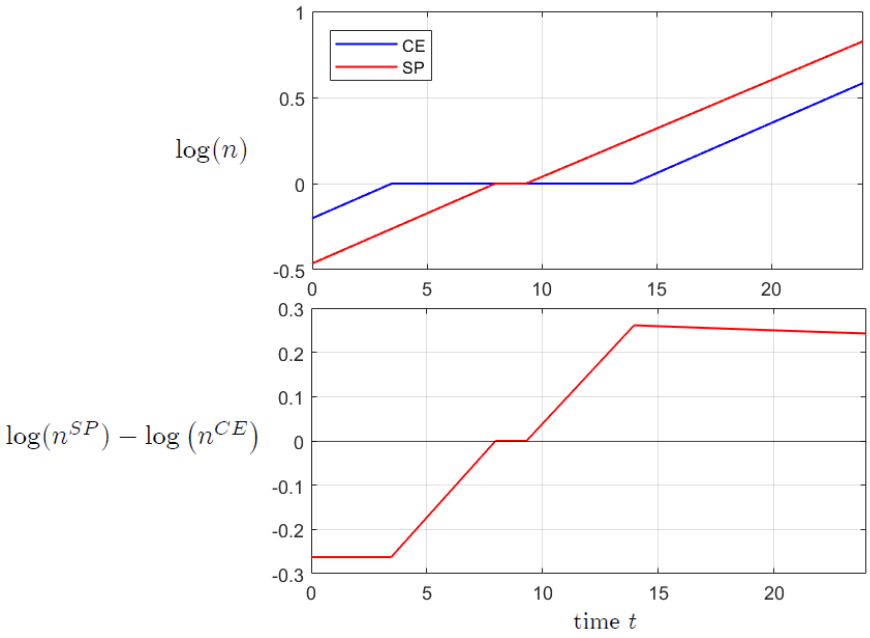


Figure 5: Comparison between CE and SP with the same initial state $(K(0), J(0))$

Wedges between CE and SP. We show in the proof that there are only two differences in the systems of equations that solve CE and SP, which result in the difference in the consumption growth path. Specifically,

	CE	SP
Stage 1:	$\frac{\dot{n}}{n} = \frac{z-\delta-\rho}{\sigma}$	$\frac{\dot{n}}{n} = \frac{z-\delta-\rho}{\sigma}$
Stage 2:	$\frac{\dot{n}}{n} = 0$	$\frac{\dot{n}}{n} = 0$
	$\tau_2 - \tau_1 = \frac{-\log(1-\lambda)}{z-\delta-\rho}$	$\tau_2 - \tau_1 = \frac{-\log(1-\lambda)}{z-\delta-\rho} + \frac{1}{z-\delta-\rho}\log \frac{n(\tau_1)^{-\sigma}/z}{n(\tau_2)^{-\sigma}/z + \underbrace{\phi\alpha n(\tau_2)^{\alpha-1}\mu(\tau_2)/z + \frac{\phi n(\tau_2)^{\alpha}}{\dot{k}(\tau_2)}\mu(\tau_2)}_{\text{future externality } (+)}}$
Stage 3	$\frac{\dot{n}}{n} = \frac{z-\delta-\rho}{\sigma}$	$\frac{\dot{n}}{n} = \frac{z-\delta-\rho}{\sigma} + \frac{1}{\sigma}\underbrace{\frac{d\log\left(1+\phi\alpha n^{\alpha-1}\mu/n^{-\sigma}\right)}{dt}}_{\text{growth of externality } (-)}$

where μ is the costate variable (i.e., the marginal valuation of J), which now has an explicit expression given in the Appendix. Note that the consumption growth path is not only determined by the Euler equation ($\frac{\dot{n}}{n}$), but also depends on the length $\tau_2 - \tau_1$ of stage 2. In fact, the second stage is a "corner solution" and the first-order condition (Euler equation) is not binding, and other

conditions are needed to determine $\tau_2 - \tau_1$ (see the proof). We discuss the economic intuition.

- *stage 3*

  In stage 3, the consumption Euler equation is binding. As shown in Section 2.3, because the planner internalizes the externality of consumption, the consumption growth rate differs between CE and SP. The difference is a function of the growth of the externality, $d \log \left(1 + \phi \alpha n^{\alpha-1} \mu / n^{-\sigma}\right) / dt$. We explicitly calculate the growth of the externality, which is negative.

- *stage 2*

  In stage 2, the equilibrium path is a "corner solution" and the consumption Euler equation is not binding. The planner internalizes the *future* positive externality of consumption on technology. Specifically, since the planner internalizes the positive externality in stage 3, the marginal valuation of capital at τ_2^- is higher for SP than for CE (i.e., $\frac{m_{\text{SP}}}{m_{\text{CE}}} > 1$),[20] which results in $\tau_2 - \tau_1$ being shorter in SP than in CE. In fact, $\tau_2 - \tau_1$ is given by $\tau_2 - \tau_1 = \frac{1}{z-\delta-\rho} \log \frac{m(\tau_1)}{m^-(\tau_2)}$ (see the proof in the Appendix). We calculate that $\tau_2 - \tau_1$ in SP is shorter than in CE by

$$(\tau_2 - \tau_1)|_{\text{CE}} - (\tau_2 - \tau_1)|_{\text{SP}} = \frac{1}{z - \delta - \rho} \log \frac{1 + \frac{z\lambda}{z-\delta} \frac{\phi n (\tau_2)^\alpha}{\dot{k}^{\text{SP}}(\tau_2)}}{1 - \frac{\lambda}{z-\delta} \phi \alpha n \left(\tau_2\right)^{\alpha-1}}.$$

- *stage 1*

  In stage 1, the consumption growth rates for CE and SP are the same. But their consumption levels differ because agents respond to the future equilibrium path.

Policy Instruments to Implement SP Allocation. The planner's allocation can be implemented in a number of ways (see, e.g., Lucas and Stokey (1983)). In our model, based on (21) and (22), policy instruments can be imposed either on the savings rate of return $r(t) = z - \delta$ or on the marginal cost of consumption. We first consider the former instrument. Suppose the government provides a subsidy (positive or negative) to the savings rate of households in CE at rate $\mathcal{S}(t)$. Then the budget constraint in the program (19) for the decentralized economy becomes

$$\dot{k}(t) = r(t)(1 + \mathcal{S}(t)) k(t) - z\mathbf{e}(n(t); J(t)) - T(t), \tag{27}$$

where $r(t) = z - \delta$ and $T(t) = r(t) \mathcal{S}(t) k(t)$ is a lump-sum tax (transfer) to households at time t. The government budget constraint is balanced period by period as in Itskhoki and Moll (2019). Proposition 2 follows.

[20] As will become clear in (A.5), $m_{\text{CE}} = \frac{\partial V_3}{\partial k}$ while $m_{\text{SP}} = \frac{\partial V_3}{\partial k} + \Lambda_K \frac{\partial V_3}{\partial \Lambda}$ in the third stage.

Proposition 2 ***(Two-phase policy intervention)*** *Suppose at $t=0$ the government can commit to a policy path given in (27). For the decentralized economy, there exists a unique optimal policy, which is two-phase, to implement the planner's solution path, that is,*

$$\mathcal{S}(t)=\begin{cases}\chi_1 & \text{for } t\in\left[\tau_1^{SP},\tau_2^{SP}\right)\\ \chi_2(t) & \text{for } t\in\left[\tau_2^{SP},\infty\right)\end{cases}$$

where $\chi_1=\frac{z-\delta-\rho}{z-\delta}\frac{-\log\vartheta}{\log\vartheta-\log(1-\lambda)}>0$ *with* $\vartheta=\frac{1-\frac{\lambda}{z-\delta}\phi\alpha n^{SP}(\tau_2)^{\alpha-1}}{1+\frac{\phi n^{SP}(\tau_2)^{\alpha}}{k^{SP}(\tau_2)}\frac{z\lambda}{z-\delta}}$, $\chi_2(t)=\frac{1}{z-\delta}\frac{\frac{\lambda}{z-\delta}\phi\alpha(\alpha-1)n^{SP}(t)^{\alpha-2}}{1-\frac{\lambda}{z-\delta}\phi\alpha n^{SP}(t)^{\alpha-1}}\frac{dn^{SP}(t)}{dt}<0$*, and* τ_1^{SP} *and* τ_2^{SP} *are the* τ_1 *and* τ_2 *in the planner's optimal path, respectively.*

Proof: *See Appendix.*

Proposition 2 establishes that the unique optimal policy intervention is a two-phase intervention with *opposite signs* at different phases. The early-phase suppression of consumption is to internalize the positive, forward-looking growth of the externality, while the later-phase boosting of consumption is to internalize the negative, forward-looking growth of the externality. From Section 2.3, it is clear how to implement SP in Stage 3. But how can a shorter $\tau_2-\tau_1$ in SP than in CE be implemented for Stage 2? Note that for CE, $\tau_2-\tau_1=\frac{1}{r(t)-\rho}\log\frac{m(\tau_1)}{m(\tau_2)}$, where $r(t)=z-\delta$. Hence, a positive subsidy rate $\mathcal{S}(t)$ on $r(t)$ would make $\tau_2-\tau_1$ shorter (see the proof).

It is important to emphasize that the interventions in both phases are indispensable for implementing the SP allocation. In Section 4.2, we quantify the effect of each intervention and show they are complementary: combining an early savings push with a later consumption subsidy delivers higher welfare than either phase alone. Intuitively, the early phase depends on the late phase because, without a subsequent "demand pull", accumulated capital cannot be fully utilized, leading to oversaving; conversely, the late phase depends on the early phase because a consumption stimulus is most effective when sufficient capital is in place to leverage technology improvements.

Alternative Implementation Through Industrial Policies. Implementation can also be through altering the marginal cost of consumption. Consider a policy that subsidizes (taxes) the capital cost of consumption-goods producers. Specifically, the subsidy rate on the capital cost of firm j is $s(j,t)$; that is, firm j only needs to pay the cost $(1-s(j,t))\varphi(j|J(t))$ when it incurs the unit cost $\varphi(j|J(t))$. Without policy, the sector's cost is $\mathbf{e}(n(t);J(t))=\int_0^{n(t)}\varphi(j|J(t))\,dj$; with policy, it becomes $\mathbf{e}(n(t);J(t))-\Psi(n(t))$, where $\Psi(n(t))=\int_0^{n(t)}s(j,t)\varphi(j|J(t))\,dj$ is the total subsidy to the sector.

The decentralized household's budget constraint in the program (19) becomes

$$\dot{k}(t)=(z-\delta)k(t)-z\left(\mathbf{e}(n(t);J(t))-\Psi(n(t))\right)-zT(t), \tag{28}$$

where $T(t)=\Psi(n(t))$ is the lump-sum tax on consumption goods producers so the government's

period budget balances. Under the policy, the CE Hamiltonian first-order condition with respect to n is

$$n^{-\sigma} - mz\left(\mathbf{e}_n\left(n\left(t\right);J\left(t\right)\right) - \Psi^{'}\left(n\left(t\right)\right)\right) = 0.$$

To implement the SP allocation, the CE and SP Euler conditions must coincide, which requires their first-order conditions to coincide. From Section 2.3, the SP first-order condition is $n^{-\sigma} - m_{\mathrm{SP}} z\left(\mathbf{e}_n\left(n;J\right) - \Upsilon\left(n,J\right)\right) = 0$, where $\Upsilon\left(n,J\right) \equiv \frac{\mu \mathcal{J}_n(n,J)}{m_{\mathrm{SP}} z}$ is the externality wedge. For Stage 3 (where the FOC is binding), this wedge is calculated as $\Upsilon\left(n,J\right) = \frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}$. To implement SP, therefore, a necessary and sufficient condition is $\Psi^{'}\left(n\left(t\right)\right) = \Upsilon\left(n\left(t\right),J\left(t\right)\right)$. Note that $\Psi^{'}\left(n\left(t\right)\right) = s\left(n\left(t\right),t\right)\varphi\left(n\left(t\right)|J\left(t\right)\right)$. Hence, $s\left(n\left(t\right),t\right)\varphi\left(n\left(t\right)|J\left(t\right)\right) = \Upsilon\left(n\left(t\right),J\left(t\right)\right)$; that is, the instrument must reduce the unit cost of the frontier good n by exactly $\Upsilon\left(n\left(t\right),J\left(t\right)\right)$. Corollary 2 follows.

Corollary 2 ***(Industrial policy)*** *Suppose the government can only target a range of industries with a fixed measure ϖ (due to various constraints). The planner's path can be implemented via a two-phase industrial policy that subsidizes the capital cost of a continuous range of varieties that overlaps the frontier (i.e., $j = n^{SP}\left(t\right)$ belongs to the subsidized set). It is not optimal to subsidize industries that are too premature ($j > n^{SP}\left(t\right) + \varpi$) or too mature ($j < n^{SP}\left(t\right) - \varpi$).*

Proof: *See Appendix.*

The details of the implementation are provided in the appendix. Intuitively, implementation is a dynamic problem: the policy must make the Euler margin coincide with the planner's by acting on the marginal cost at the frontier. Subsidizing varieties that are too mature is ineffective at the margin—it lowers the level $\mathbf{e}\left(n;J\right)$ via an income effect but leaves the marginal cost at the frontier unchanged. Subsidizing varieties that are too premature has no contemporaneous demand and thus no effect at the Euler margin. Only a range that overlaps the frontier directly reduces the marginal cost, thereby delivering the planner's wedge under the policy constraint.

The rationale for industrial policy in our model is not the removal of static distortions but the correction of dynamic externalities captured by the wedge between the CE and SP Euler conditions. Targeting emerging industries is therefore necessary. This echoes Rodrik (2004), who argues that government policy should target new industries. We provide a dynamic rationale for that emphasis.

3 Externalities and Spillovers in the Investment Goods Sector

This section serves two main purposes. First, we incorporate learning-by-doing externalities in the investment goods sector and show model robustness. Second, we demonstrate potential development traps. So far the learning-by-doing mechanism affects the growth in the transition but not in the steady state. In fact, the steady-state growth rates given in Corollary 1 are all independent of

the production externality characterized by parameter λ. We now extend the model to show that steady-state growth depends on the externality forces during the transition.

In modeling externalities in physical capital, Romer (1986) assumes that an economy's knowledge stock increases with its capital stock. On the other hand, technological accumulation in one sector often has spillover effects on other sectors (e.g., Jaffe (1986), Aghion and Jaravel (2015), Liu and Ma (2022)). In our model context, it is natural that technological accumulation in the consumption goods sector spills over into the investment goods sector. To capture both effects, we assume that productivity in the investment goods sector is

$$z(t)=\begin{cases}\mathcal{Z}\left(K(t),z\psi\left(J(t)\right)\right) & & \text{for } t\leq t^{*}\\ \mathcal{Z}\left(K(t),z\psi\left(J(t)\right)\right) & \text{if } J(t^{*})<\bar{J} & \text{for } t>t^{*}\\ \mathcal{Z}\left(K(t),z^{H}\psi\left(J(t)\right)\right) & \text{if } J(t^{*})\geq\bar{J} & \text{for } t>t^{*}\end{cases}, \tag{29}$$

where $\mathcal{Z}(x,y)=\left[\iota x^{\frac{\varsigma-1}{\varsigma}}+(1-\iota)\,y^{\frac{\varsigma-1}{\varsigma}}\right]^{\frac{\varsigma}{\varsigma-1}}$ with $\varsigma\in[0,1)$, and $\psi(\cdot)$ is a continuous function satisfying $\psi(0)=0$, $\psi^{'}>0$ and $\lim_{J\to\infty}\psi(J)=1$, and t^{*}, $\bar{J}$, z, and z^{H} are parameters with $z^{H}>z$. Equation (29) states that productivity in the investment goods sector depends on capital stock $K(t)$ as well as the technology frontier (experience capital) $J(t)$ in the consumption goods sector, which exhibit *gross complementarity* in CES. In particular, if and only if the technology frontier of the consumption goods sector exceeds the threshold $\bar{J}$ by time point t^{*}, the technology in the investment goods sector can upgrade to a higher level, from z to z^{H}. The above setting is in the similar spirit of the big push: a faster development in the consumption goods sector leads to a permanent increase in efficiency in the investment goods sector. This permanent increase results in a higher long-run growth rate.

Investment Goods Sector Externalities. To facilitate the analysis, we first focus on the case in which z^{H} is not realized in equilibrium (i.e., the condition $J(t^{*})<\bar{J}$ holds true in the equilibrium path). In this case, (29) is simplified to $z(t)=\mathcal{Z}\left(K(t),z\psi\left(J(t)\right)\right)$; without causing confusion, we write function $\mathcal{Z}$ as $\mathcal{Z}(K,J)$. Following the analysis in Section 2.3, we find that the subsidy to the savings rate of return in CE at rate $\mathcal{S}(t)$ can implement the SP's allocation, where

$$\mathcal{S}(t)=\frac{1}{\mathcal{Z}(K,J)-\delta}\left(\underbrace{\mathcal{Z}_{K}(K,J)\left(k-\mathbf{e}(n;J)\right)}_{\textbf{investment externality}}+\underbrace{\frac{d\log\left(1+\mathcal{J}_{n}(n,J)\mu/n^{-\sigma}\right)}{dt}}_{\textbf{growth of consumption externality}}\right). \tag{30}$$

Comparing (30) with (23), we see a new term arising due to investment externalities. This is consistent with Romer's (1986) model implication that capital externalities cause the social planner to want to subsidize savings at any given point in time. This new force strengthens our result regarding the early-phase policy of suppressing consumption. Importantly, because $\mathcal{Z}_{K}(K,J)\to 0$

as $\frac{K}{J} \to \infty$ (i.e., CES gross complementarity between K and J under $\varsigma < 1$), the new force diminishes over time, given that $\frac{K(t)}{J(t)}$ increases over time in the equilibrium dynamics. Therefore, this new force will not qualitatively alter our result regarding the late-phase policy of stimulating consumption, because $K(t)$ and $J(t)$ are complementary in improving $z(t)$.

Development Traps. If the consumption goods sector cannot grow fast enough during transition, the productivity of the investment goods sector cannot improve and upgrade to a higher level and is possibly permanently trapped at a lower level, which would result in a lower steady-state growth rate. Now we consider the case in which z^H is potentially realized in equilibrium. To simplify the analysis, we focus on $\varsigma \to 0$, under which the function $\mathcal{Z}$ becomes a Leontief function. We also focus on t^* being high enough so that, in equilibrium, $\frac{K(t)}{J(t)}$ at $t = t^*$ is high enough such that $\frac{K(t)}{z^H \psi(J(t))} > 1$ at $t = t^*$. Under such parameter values, (29) becomes

$$z(t) = \begin{cases} \min\left(K(t), z\psi\left(J(t)\right)\right) & \text{for } t \leq t^* \\ z\psi\left(J(t)\right) & \text{if } J(t^*) < \bar{J} \quad \text{for } t > t^* \\ z^H \psi\left(J(t)\right) & \text{if } J(t^*) \geq \bar{J} \quad \text{for } t > t^* \end{cases}. \tag{31}$$

Hence, after $t = t^*$, $z(t)$ is either $z\psi\left(J(t)\right)$ or $z^H\psi\left(J(t)\right)$. Recalling that $\lim_{J\to\infty} \psi(J) = 1$, it is easy to show that under (31), the steady-state growth rate of any aggregate variable converges to one of the two values: the lower steady state and the higher steady state. For example, for the GDP growth, the steady state growth rate is one of the following two values:

$$\lim_{t\to\infty} g_N(t) = \lim_{t\to\infty} g_K(t) = \frac{(z-\delta)-\rho}{\sigma} \text{ or } \frac{\left(z^H-\delta\right)-\rho}{\sigma}.$$

Crucially, even a small difference between two paths for $t < t^*$ can result in two different steady-state growth rates. In comparing the CE and the planner's solution, the planner can achieve a higher $J(t)$ at $t = t^*$ than the CE. Therefore, for some threshold $\bar{J}$, only the planner's solution can achieve the higher steady state. Proposition 3 follows.

Proposition 3 *Government intervention in savings during the transition can achieve the higher steady-state growth that the laissez-faire equilibrium cannot.*

***Proof**: See Appendix.*

Our model does not feature multiple equilibria; rather, the presence of the externality justifies government intervention to achieve the higher steady-state growth. This result echoes the quantitative finding of Buera et al. (2021) that even without equilibrium multiplicity, the elements responsible for the complementarity substantially amplify the effect of distortions, justifying policy intervention.

4 Empirical Evidence and Quantitative Analysis

In this section, we provide empirical evidence to support the model and quantitatively evaluate the model's implications.

4.1 Empirical Evidence

4.1.1 Historical Evidence on Consumption Upgrading

Figure 6 illustrates the chronology of the commercialization of some consumer products in the United States, South Korea, and China. It clearly shows a pattern of consumption upgrading: initially, household consumption bundles consist primarily of basic goods but later shift towards more advanced products. For example, prior to World War II, American households typically consumed items such as sewing machines and bicycles, subsequently upgrading to more sophisticated goods such as color televisions, refrigerators, and air conditioners. A similar pattern is evident in other countries, including South Korea and China.

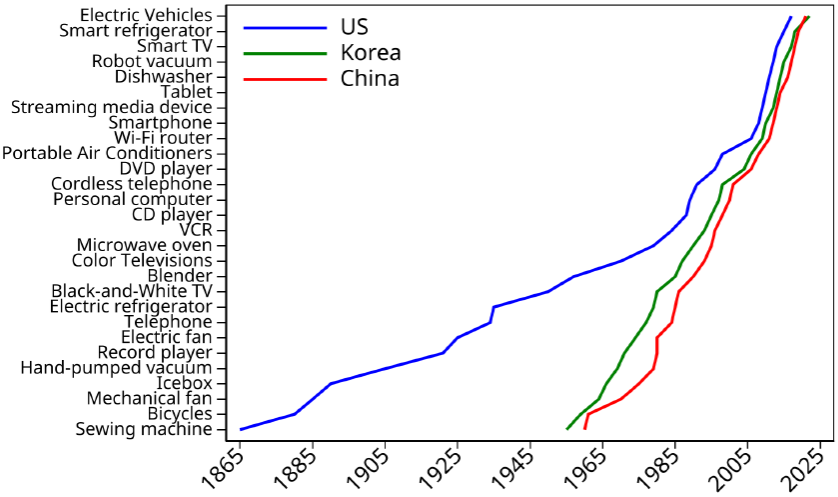


Figure 6: Chronology of commercialization of some consumer products (source: ChatGPT)

Note: The horizontal axis is the year of commercialization of a consumption item.

4.1.2 Test on the Model

Hypothesis. The model's core mechanism posits that learning-by-doing externalities and spillovers induced by consumption follow a hump-shaped trajectory over the development process: initially weak when consumption concentrates on basic goods below the technology frontier, then intensifying as consumption crosses the technological frontier, and diminishing in later stages when the economy gets sufficiently close to the world technology frontier. While direct empirical verification of this

dynamic pattern requires rich longitudinal data that are rarely available across countries, we test a *sufficient condition* underpinning this mechanism: the production of advanced goods generates significantly stronger production externalities than basic goods. Specifically, we test the hypothesis:

Hypothesis: Market size for advanced goods has a significantly larger positive impact on productivity growth than market size for basic goods.

In the context of our model with hierarchical-demand preferences, this hypothesis is consistent with the key assumption $\phi > \phi_0$ in Assumption 3 of our model.

Empirical Strategy. Our empirical methodology follows Boppart and Weiss (2013), who estimate the effect of market size on productivity growth within the same sector to study industry-directed technical change, and Beerli, Weiss, Zilibotti, and Zweimüller (2020), who use the Chinese manufacturing industry data with good-level market size and associated industry productivity to examine demand forces of technical change, employing an identification strategy to address the endogeneity of market size. Specifically, we use cross-country dataset to construct country×year×goods level stocks of consumption durables and then merge this with the country×year×industry level productivity data. We analyze the relationship between a good's market size and the productivity growth of its associated sector, as in Beerli, Weiss, Zilibotti, and Zweimüller (2020). Crucially, we extend their methodology by distinguishing between basic and advanced goods and testing whether the market size-productivity relationship differs significantly across good types.

Datasets and Variable Construction. We use a dataset that is constructed by combining two different sources: BvD Orbis and IPUMS (International). The BvD Orbis dataset covers private and public firms worldwide. It provides standardized, detailed annual financial statement data for each firm. We use this data to calculate productivity measures at firm and industry levels. The second source is IPUMS (International), which provides information on the quantity of durable goods consumption.[21] The dataset is based on surveys conducted by national statistical agencies in various countries, which report individual-level purchases of 11 types of durable goods (air conditioner, autos, computer, heat, hot water, radio, refrigerator, television, VCR, washer).

To construct the market size of each durable good, we use the following formula:

$$\mathrm{MS}^{\text{actual}}_{i,j,t+1} = \underbrace{\left[\mathrm{stock}^{\text{actual}}_{i,j,t+1} - \mathrm{stock}^{\text{actual}}_{j,t}\right]}_{\text{new purchases}} + \underbrace{\delta_j \cdot \mathrm{stock}^{\text{actual}}_{i,j,t}}_{\text{replacement purchases}} ,$$

where the measured market size of durable good j in country i from time t to $t+1$ is the sum of new purchases and replacement purchases. The former term, $\mathrm{stock}^{\text{actual}}_{i,j,t+1} - \mathrm{stock}^{\text{actual}}_{j,t}$, represents

[21] The cross-country datasets that provide quantity information on consumption durables are rare. Most datasets contain expenditure information, which is the product of quantity and prices (such as the EUKLEMS). For a limited set of consumption durables, the IPUMS dataset provides the quantity information and therefore allows us to construct the stock of consumption durables.

the observed new purchases of durable goods from time t to $t+1$. The latter term, $\delta_j \cdot \text{stock}_{i,j,t}^{\text{actual}}$, represents the assumed replacement from time t to $t+1$, where the replacement rate δ_j is product-specific (from the Bureau of Economic Analysis (BEA)).

Note that the market size measured here is potentially endogenous: the market size of a durable good can be affected by the price or the quality of the goods, which in turn can be affected by the innovation from firms producing that good. To address this endogeneity, we follow the approach in Acemoglu and Linn (2004) and Beerli, Weiss, Zilibotti, and Zweimüller (2020) to construct a "potential" market size by using the variations from household income dynamics. The idea is that as the population of wealthy individuals increases, so does the potential market size for a durable good. Specifically, for all households in the IPUMS dataset, we sort households into four income groups: low income (below 11,900 USD), lower middle income (11,900-21,458 USD), upper middle income (21,458-35,770 USD), and high income (35,770 USD or more). We then use the U.S. households' actual purchases to estimate the probabilities of purchasing each durable at different income quartiles. These purchase probabilities are then applied to compute the potential market size for each durable good in other countries.[22]

Importantly, we sort the industries into two groups: advanced goods and basic goods. More specifically, we define basic goods as those that were massively commercialized before WWII in the U.S. and advanced goods as those that were commercialized after WWII (see Table 1). The advanced goods are generally more technologically sophisticated and include air conditioners, automobiles, computers, refrigerators, televisions, and VCRs. The basic goods are less sophisticated and include heating equipment, hot water heaters, radios, and washing machines.

Table 1: Mass consumption of household goods in the U.S. with sources

[22] Note that the IPUMS dataset does not report income distribution except for U.S. and Brazil. Therefore, we cannot construct purchase probabilities for each country.

Basic Goods		
Product	Period of Mass Consumption	Source
Radio	1930s	US Census Bureau data (1940)
Washing Machine	1940s-1950s	Life Magazine (1950)
Heating Equipment	1930s-1940s	US Housing Census (1940)
Hot Water Heater	1940s	American Builder Magazine (1947)
Advanced Goods		
Air Conditioner	1960s-1970s	New York Times (July 1970)
Automobile	1950s	US Department of Transportation data (1956)
Computer	1990s	Washington Post (1998)
Refrigerator	1950s	Consumer Reports (1955)
Television	1950s	CBS market research (1955)
VCR	1980s	Wall Street Journal (1988)

The construction of sectoral productivity is relatively standard. We first use the Levinsohn and Petrin Method to construct total factor productivity (TFP) at the firm level. We then aggregate the firm-level TFP to the industry level by taking the simple (unweighted) average within each country×sector×year cell. Finally, we merge the market size data with the TFP data based on the concordance shown in Table 2. Eventually, we get a year×sector×country panel containing information on both the market size and the productivity of the associated industry.

Table 2: Mapping consumer durables to sectors

Product Name	NAICS Code	Industry Name
advanced good		
aircon	333415	air-conditioning and industrial refrigeration
autos	336111	automobile
computer	334111	electronic computer
refrig	335220	major household appliance
tv	334310	audio and video equipment
vcr	334310	audio and video equipment
basic good		
heat	333414	heating equipment (except warm air furnaces)
hotwater	335220	major household appliance
radio	334310	audio and video equipment
washer	335220	major household appliance

Note: This table shows the mapping between the consumption durables and their corresponding NAICS (North American Industry Classification System) codes. Products are grouped into advanced and basic consumer durables.

Empirical Results. To test the asymmetric effect of market size on industry productivity, we

run the following simple panel regression:

$$\text{Productivity}_{i,j,t} = \alpha_0 + \beta \cdot \text{MS}_{i,j,t} + \gamma \cdot \text{Advanced}_j + \delta \cdot (\text{MS}_{i,j,t} \times \text{Advanced}_j) + \tau_t + \kappa_{i,s} + \epsilon_{i,j,t},$$

$\text{Productivity}_{i,j,t}$ (in logs) is the average log(TFP) for the associated sector. $\text{MS}_{i,j,t}$ is the market size (in logs) of durable good j in country i at time t, Advanced_j is a dummy variable that equals 1 if the durable good j is an advanced good, and $\kappa_{i,s}$ are country-sector fixed effects and τ_t is the year fixed effect.

The results are shown in Table 3. The first column shows the regression coefficient for all the durable goods. The coefficient is overall positive but not statistically significant, suggesting that the market size of a specific durable good has a weak positive impact on productivity growth. However, the second column shows the results when the sample is restricted to advanced goods. The coefficient of the interaction term between market size and the advanced good dummy is 0.47 with a p-value of less than 0.01, suggesting that the impact of market size on productivity growth is significantly larger for advanced goods than for basic goods. The third column shows the results when we use the potential market size instead of the realized market size. While there is no statistically significant relationship between market size and productivity growth for all goods, the coefficient of the interaction term between market size and the advanced good dummy is 2.48 with a p-value of 0.002, suggesting that the relationship between market size and productivity growth is robust to the endogeneity of market size.

Table 3: Market size and productivity

	(1)	(2)	(3)	(4)
Market Size (Realized)	0.0407	0.0255		
	(0.1002)	(0.1696)		
Advanced Goods × Market Size (Realized)		0.4692***		
		(0.0029)		
Market Size (Potential)			-0.0893	-0.6527
			(0.7977)	(0.1455)
Advanced Goods × Market Size (Potential)				2.4756***
				(0.0020)
Year FE	yes	yes	yes	yes
Country-Good FE	yes	yes	yes	yes
Observations	124	124	149	149
R^2	0.8363	0.8493	0.8026	0.8135

Note: The dependent variable in all columns is log(TFP). Column (1) shows the effect of realized market size. Column (2) adds an interaction with the Advanced Goods indicator. Column (3) shows the effect of potential market size. Column (4) adds an interaction with the Advanced Goods indicator. All specifications

include year and country-good fixed effects. Standard errors are clustered at the country-good level. p-values in parentheses. Significance levels: *** p < 0.01, ** p < 0.05, * p < 0.10.

Chinese Evidence. Beerli, Weiss, Zilibotti, and Zweimüller (2020) use Chinese data to study demand forces of technical change. The advantage of using Chinese data is that household survey data on the stock of consumption durables are available. We now use the same datasets[23] and follow the same empirical strategy as theirs, but distinguish between basic and relatively advanced goods to test our hypothesis. Using the Chinese dataset, we are able to get more types of consumption durables: air conditioner, camera, car, computer, electric fan, homevideo appliances, kitchen appliances, motorcycle, radio, refrigerator, satellite dish, sewing machine, telephone, washing machine, and bicycle.

In the case of China, we define basic goods and advanced goods in the following way. If the realized market size of a good decreased during the observation period (1998-2007), we define it as a basic good; otherwise, we define it as an advanced good. Of these, bicycle, electric fan, home video appliances, motorcycle, sewing machine, and telephone are considered basic durable goods. Consistent with the international evidence, these goods are less sophisticated.

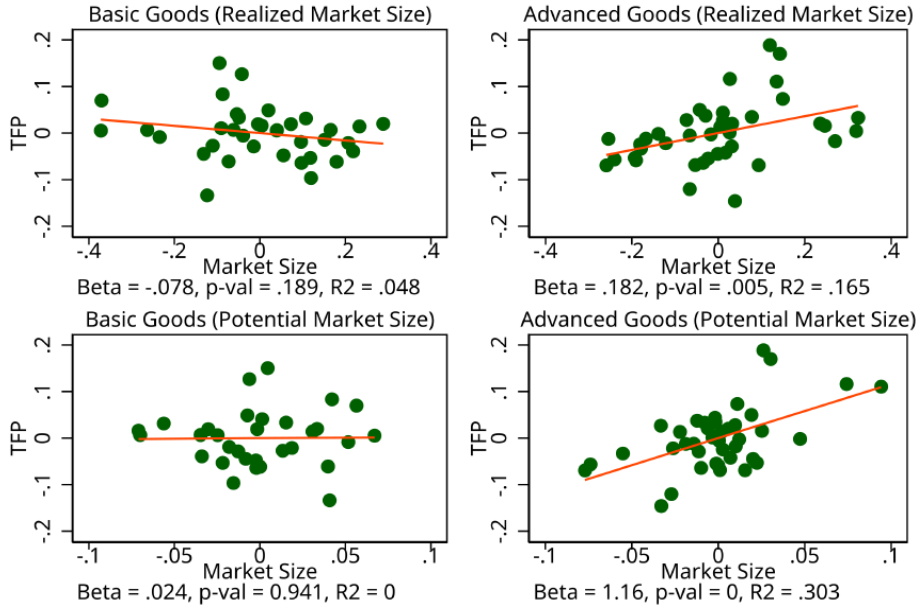


Figure 7: Market size and productivity for basic goods and advanced goods

Note: The figure shows the residual scatter plots by controlling the industry and year fixed effect. The

[23] Specifically, for firm level datasets, we use the Annual Survey of Industrial Production (ASIP) dataset (1998 to 2007), which is a widely used census of all state-owned and non-state-owned industrial firms with annual sales above 5 million RMB. We use this dataset to construct industry level productivity measures. For the household level dataset, we use the China Family Panel Studies (CFPS) dataset, which is a nationally representative longitudinal survey conducted by Peking University. The CFPS collects detailed information on household demographics, income, consumption, and ownership of durable goods. The dataset covers multiple waves from 1998 to 2007, allowing us to track changes in durable goods ownership over time. We use this dataset to construct the realized and potential market size of consumption durables.

horizontal axis is the market size of a specific durable good, and the vertical axis is the simple (unweighted) average log(TFP) in that industry. Each dot represents a year×good observation.

Figure 7 shows the residual scatter plots between the market size of these basic goods and the advanced goods (after controlling for sector and year fixed effects). For basic goods, the regression slope is largely muted: -0.08 for the realized market size and 0.02 for the potential market size (neither is statistically significant). By contrast, for advanced goods, the regression slope is positive and highly significant: 0.18 for the realized market size and 1.16 for the potential market size. This is consistent with the cross-country result, where the market size impact is asymmetric between basic and advanced goods.

4.2 Quantitative Analysis

In this subsection, we quantify the effects of the two-sided policies.

Model Parameterization. To parameterize the model, we externally assign a group of parameters and internally calibrate the remaining. For the assigned ones, we set the depreciation rate of physical capital, δ, to 0.07, following Greenwood, Hercowitz, and Krusell (2000). We set the discount rate $\rho = 0.04$ and the intertemporal elasticity of substitution $1/\sigma = 1/1.5$, values which are in line with recent quantitative work (e.g., Boppart and Krusell (2020)). We then choose the cost reduction parameter λ to be 0.6, following Buera et al. (2021), and the curvature of the learning-by-doing externality parameter α to be 0.8, based on the estimation in Halter, Oechslin, and Zweimüller (2014). Finally, we normalize b (the natural growth rate of J) to 0, J_0 (the initial value of J) to 1, and ε (the curvature of the unit cost function) to 0—approximated by a very small positive value in the simulation to ensure a smooth function suitable for numerical simulation.[24] We also normalize $\phi_0 = 0$ to reflect the stylized fact that the market size for basic goods has minimal or no impact on productivity.

The remaining four parameters are internally calibrated: ϕ (the efficiency gap of experience production between advanced and basic goods), z (the raw return on capital), K_0 (the initial value of capital), and ζ (the experience accumulation curvature).[25]

We use the model to generate six moments that capture the growth and transition dynamics of an economy. The first moment is the starting time of the rise of consumption (akin to τ_2 in the baseline model), defined as the point at which $n(t)$ takes off from $J(t)$ with $n(t) - J(t)$ exceeding 0.001 for the first time; denote this point in time by $t = \tau^*$. The second set of moments are the average growth rates of $N(t)$ and $K(t)$ over the period $t \in [0, \tau^*)$. The third set of moments are the maximum growth rates of $N(t)$ and $K(t)$ over the lifetime path $t \in [0, +\infty)$. The final moment

[24] When $\varepsilon \to 0$, the parameter κ plays no role.

[25] In the quantitative analysis, we choose ν to normalizes $\frac{\nu}{\zeta} = 1$, meaning that the function $\beta(N, J)$ takes the standard logistic form as in the technology diffusion literature.

is the impact of market size on the productivity for advanced goods.[26]

Table 4: Model parameterization

Parameter	Symbol	Value	Source/Notes
externally assigned parameters			
capital depreciation	δ	0.070	Greenwood et al. (2000)
discount rate	ρ	0.040	–
(inverse) of elasticity	σ	1.500	Boppart and Krusell (2020)
efficiency gap	λ	0.600	Buera et al. (2021)
return to scale of J	α	0.800	Halter et al. (2014)
natural growth of J	b	0.000	normalized
initial technology	J_0	1.000	normalized
experience accumulation (early)	ϕ_0	0.000	normalized
internally calibrated parameters			
experience accumulation (later)	ϕ	0.0581	–
capital accumulation	z	0.1973	–
initial capital	K_0	0.7052	–
experience accumulation (curvature)	ζ	0.0749	–

Our choice of six data moments is as follows. For the length of the transition period before τ^*, we follow Matsuyama (2002) and set it to 25 years.[27] For the capital growth rate, we follow Barro (2003) to use 5.7% as the average growth rate during the transition period and 7.0% as the maximum growth rate. For the data counterpart of N_t, we follow Greenwood and Uysal (2005), who use the number of new trademarks to measure the variety of consumption goods. They show that the number of new trademarks in the U.S. increased sixfold from 1940 to 1990, which is equivalent to an average growth rate of 3.6%. We use this number as the average growth rate of $N(t)$ in the transition. New trademarks tripled from 1970 to 1990, implying a maximum growth rate of $N(t)$ of 5.7%. For the market size impact on productivity for advanced goods, we use the OLS estimate from our empirical analysis, which is 0.47.[28]

We calibrate the model by matching theoretical moments and data moments. Denoting the vector of the four endogenously estimated parameters by $\Theta = (z, \phi, K_0, \zeta)$, we choose the parameters to minimize the loss function: $L(\Theta) = \sum_{i=1}^{6} \left(\frac{\text{moment}_i^{\text{model}}(\Theta) - \text{moment}_i^{\text{data}}}{\text{moment}_i^{\text{model}}(\Theta) + \text{moment}_i^{\text{data}}} \right)^2$, where $\text{moment}_i^{\text{model}}$ is the value of moment i in the model and $\text{moment}_i^{\text{data}}$ is the value of the moment in the data.

[26] This corresponds to $\frac{d \log\left(\frac{1}{\varphi(j=J|J)}\right)}{d \log n} = \frac{dJ}{d \log n} \frac{-d(\log \varphi(j=J|J))}{dJ}$, where the TFP of the production of consumption varieties is $\frac{1}{\varphi(j|J)}$ and advanced goods are proxied by emerging goods ($j = J$). We use the maximum value over the entire simulation path to match the empirical counterpart, meaning the welfare gain calculated through the calibrated model is a lower bound.

[27] Alternatively, we can follow Buera and Shin (2013) to set it to 30 years. The main quantitative results are similar.

[28] The reason that we use the OLS coefficient is that, in the model, market size is endogenous. Therefore, using the OLS coefficient from the data counterpart is more consistent with the model.

Although the four model parameters are jointly estimated, some moments are more informative for certain groups of parameters. For instance, the average and maximum growth rates of $N(t)$ and $K(t)$ are more informative for z. The transition period is more informative for K_0. The market size impact on productivity is more informative for ϕ. Table 5 reports the moment matching results.

Table 5: Moments

Moment	Model	Data	Data Source
years to rise of consumption	27.399	25.000	Matsuyama (2002)
market size impact on productivity	0.528	0.470	our regression result
growth of K (max)	0.069	0.070	Barro (2003)
growth of K (average)	0.048	0.057	Barro (2003)
growth of N (max)	0.058	0.057	Greenwood and Uysal (2005)
growth of N (average)	0.037	0.036	Greenwood and Uysal (2005)

Policy and Quantitative Implications. We use the calibrated model to assess the effects of the two-sided policy. Specifically, the optimally designed two-sided policy is $\mathcal{S}(t)$, as given in (23). Under this policy, there exists a unique turning point in time, before which the subsidy rate is positive and after which it is negative. We consider two alternative policies for counterfactuals: a one-sided, early-stage truncated policy (i.e., $\max\{\mathcal{S}(t), 0\}$) and a one-sided, late-stage truncated policy (i.e., $\min\{\mathcal{S}(t), 0\}$).

Table 6: Effects of policies

Policy Type	Welfare Gains (%)	Years to Double J_0
calibrated model		
optimal policy(two-sided)	1.342	24.37
one-sided policy only (early)	0.235	25.12
one-sided policy only (late)	-1.161	27.78
no policy	0.000	28.39
closed-form solution model		
optimal policy (two-sided)	4.003	22.05
one-sided policy only (early)	2.942	22.38
one-sided policy only (late)	0.209	26.37
no policy	0.000	26.59

Note: Welfare gains in consumption-equivalent terms (i.e., % increase in consumption in every period). Years to double J_0 is the time from $t = 0$ to $J(t) = 2J_0$.

Table 6 reports the results. Column 1 shows that the consumption-equivalent welfare gain of the two-sided policy is 1.342%. This magnitude is in line with the macro development literature (e.g., 0.503%-1.533% in Itskhoki and Moll (2019) on financial frictions) and roughly one order of magnitude higher than the welfare gain attributed solely to externalities in the business-cycle

literature (e.g., 0.15% in Boissay, Collard, and Smets (2016)). Note that our model features no friction; the externality is the sole source of the policy's welfare gains. Column 2 reports the effects of different policies on the technological progress $J(t)$. The two-sided policy results in the fastest transition, reducing the time taken to double the initial J to 24.37 from 28.39 (no policy).

The bottom panel of Table 6 also reports the results for the closed-form model in Section 2.4, using the same parameter values as the calibrated model, except taking $\varepsilon \to 0^+$ and $\zeta \to 0^+$. This limiting case with kinks yields an upper-bound welfare gain of 4.003%. This case also helps to intuitively and transparently illustrate why the welfare gain in our model is significant. In fact, we find that $\tau_2 - \tau_1$ equals 10.50 in CE and 1.32 in SP in Figure 5. In other words, the government would seek to engineer a 'demand pull' to accelerate industrialization by shortening the intermediate no-externality period by 9.18 years, which explains the significant welfare gain in our model.

The early- and late-stage policies are interdependent. Implementing only the later-stage consumption subsidy, without an early-stage savings push, would in fact lead to a welfare loss of 1.161%. A consumption stimulus is effective only when enough capital is already in place to leverage technology improvements; without that capital base, the subsidy can be counterproductive. Conversely, if the later-stage consumption subsidy is not implemented, the welfare gain would fall to 0.235%. This raises a design question: if the government is unable to implement a late-stage subsidy (for various reasons), how should it adjust the early-stage intervention? Formally, we solve the following optimal one-sided early-stage policy problem: given that intervention takes the form of *intensity*$\times \max\{\mathcal{S}(t), 0\}$, what is the optimal intensity? Figure 7 shows that welfare is hump-shaped in intensity, peaking around 0.6. This means that without the later-stage intervention, the early-stage push should be scaled back by around 40%; otherwise, a full-scale early-stage capital push would result in excessive capital accumulation. For the analytical model, the optimal intensity is about 90% of the full-scale policy.

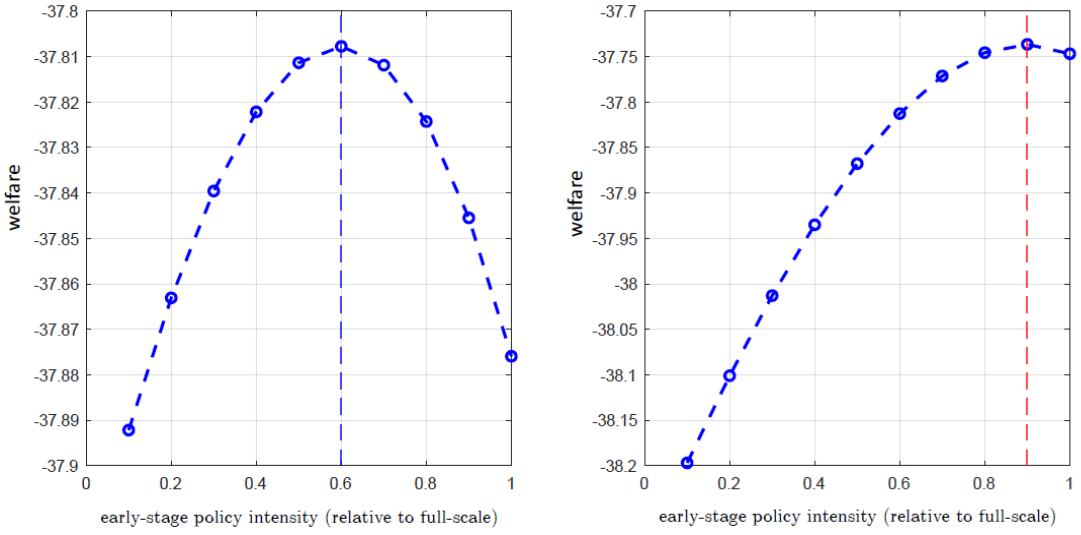


Figure 8: Optimal early-stage policy intensity if no late-stage policy (Left: for the calibrated model; Right: for the closed-form model)

5 Model Extension

We conduct several extensions on the main model, to show robustness and provide additional insights.

5.1 Microfoundation for the Technology Process

We provide a microfoundation for the law of motion of $J(t)$, as specified in Assumption 3. A key aspect is the asymmetry property ($\phi > \phi_0$). The mechanism might be that producing goods with immature technology naturally generates more knowledge (experience) as a by-product. Alternatively, firms may be incentivized to invest in research and learning when producing with immature technology, as in Romer (1990) and Aghion and Howitt (1992). Here, we model the former.

We interpret J as the economy-wide experience capital as in Dasgupta and Stiglitz (1988). The accumulation of J follows this process:

$$\dot{J} = bJ + \int_0^{\infty} j^{-(1-\alpha)} f(j, J)\, Y(j, t)\, dj, \tag{32}$$

where the second term on the RHS represents the aggregate new knowledge generated at time t from production in all industries, and $Y(j,t)$ denotes the total output of industry j at time t, which equals 1 for $j \leq N(t)$ or 0 for $j > N(t)$ in equilibrium. The two additional terms, $j^{-(1-\alpha)}$ and $f(j, J)$, act as weights. The term $j^{-(1-\alpha)}$, which decreases with j, captures the notion that experience in more technologically sophisticated industries has a lower weighting (probably because experience becomes more specific and harder to be generally applicable as j increases). The term $f(j, J)$ captures the property that the weighting also depends on the relative position of j and J, reflecting the combination of two natures: 1) *Asymmetry.* For a given J, $f(j, J)$ is higher for $j > J$ than for $j < J$, reflecting that producing goods with immature technology generates more experience than producing goods with mature technology (e.g., the evidence in Lamoreaux, Raff, and Temin (1998)). 2) *Distance.* The greater the distance between j and J, the lower the level of experience generated.

Concretely, we specify $f(j, J)$ as

$$f(j, J) = \alpha\phi_0 + \frac{\phi - \phi_0}{1+u}\left(\alpha + \frac{u}{1+u}\frac{j}{\zeta}\right) \text{ with } u = \frac{\nu}{\zeta}\exp\left(-\frac{j-J}{\zeta}\right). \tag{33}$$

In the limit $\zeta \to 0^+$, $f(j, J) = \alpha\phi_0$ if $j < J$ and $\alpha\phi$ if $j > J$. Figure 9 plots $f(j, J)$ under the

calibrated parameter values in Table 4, where we fix $J = J_0 = 1$. As can be seen in the figure, $f(j, J)$ reaches to its maximum at around $j = J$ (reflecting distance), and it is asymmetric around $j = J$.

Using (33), it follows that $\int_0^N j^{-(1-\alpha)} f(j, J)\, dj = \beta(N, J) N^{\alpha}$, where $\beta(N, J) = \phi_0 + \frac{\phi - \phi_0}{1 + \frac{\nu}{\zeta} \exp\left(-\frac{N-J}{\zeta}\right)}$. Therefore, we obtain the exact form of $\dot{J} = bJ + \beta(N, J) N^{\alpha}$ as specified in Assumption 3.

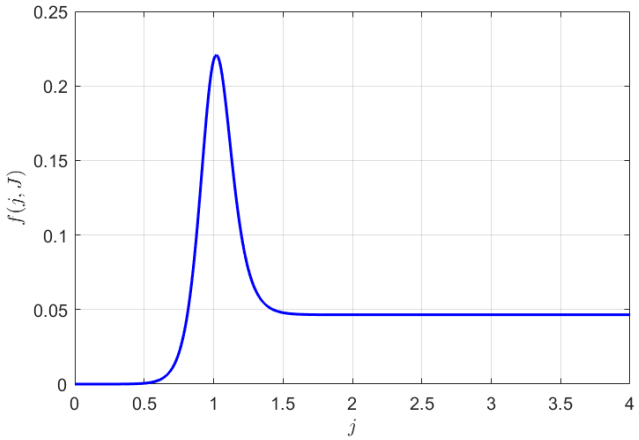


Figure 9: Weight function $f(j, J)$

5.2 Alternative Setting for Hierarchical Demand

We show our model is robust to the alternative setting for hierarchical demand that simultaneously allows for: 1) higher-indexed varieties yield higher marginal utility, in the spirit of Stokey (1988); 2) consumption choices involve an intensive margin rather than a 'zero-one' choice. Intuitively, as long as a higher-indexed variety incurs an even higher marginal cost, hierarchical demand will follow in equilibrium and our model works. The extended model that allows for an intensive margin converges to our baseline model in the limit, so both models have the same policy implications.

Households. A consumption bundle consists of a continuum of varieties indexed by j:

$$c(t) = -\frac{1}{2} \int_0^{\Delta} j^{-\gamma} (s - c)^2\, dj + \frac{1}{2} \int_{\Delta}^{\infty} j^{-\gamma} \left[s^2 - (s - c)^2 \right] dj, \tag{34}$$

where Δ is a small positive constant. The non-homothetic preference in (34) follows Bertola, Foellmi and Zweimüller (2014, Chapter 12). That is, the consumption bundle $c(t)$ is an aggregation of quadratic utility as in Hall (1978) on consumption varieties with a hierarchy factor $j^{-\gamma}$. Here, γ can be $\gamma \in (0, 1)$, in which case the hierarchy factor is a discount factor. Moreover, γ can also be negative, in which case the hierarchy factor increases with j in the spirit of Stokey (1988). For our purpose, we also allow for $\gamma > 1$.[29]

[29] To ensure that the utility function does not diverge when $\gamma > 1$, the first term $\int_0^{\Delta} j^{-\gamma} (s - c)^2\, dj$ is introduced; see Bertola, Foellmi and Zweimüller (2014, Chapter 12).

Production Function for Consumption Goods. The unit cost function is changed to

$$\hat{\varphi}\left(j|J\left(t\right)\right)=j^{\theta}\cdot\varphi\left(j|J\left(t\right)\right), \tag{35}$$

where $\theta>0$ and $\varphi\left(j|J\left(t\right)\right)$ is given in (6). The additional term j^{θ} captures that a higher-indexed variety incurs a higher marginal cost. As shown in the baseline model, $p\left(j,t\right)=R\left(t\right)\hat{\varphi}\left(j|J\left(t\right)\right)$.

Assumption. We assume $\gamma+\theta>0$. This means that if a higher-indexed variety has higher marginal utility ($\gamma<0$), it incurs an even higher marginal cost ($\theta>-\gamma$).

Hierarchical demand. It is easy to verify that in equilibrium hierarchical demand for consumption varieties still follows; that is, the first $n\left(t\right)$ are consumed in positive amounts. Given $n\left(t\right)$, we find the consumption quantity for any variety $j\in\left(0,n\left(t\right)\right]$:

$$c\left(j,t\right)=s\left(1-\frac{\varphi\left(j|J\left(t\right)\right)j^{\gamma+\theta}}{\underbrace{\varphi\left(n\left(t\right)|J\left(t\right)\right)n\left(t\right)^{\gamma+\theta}}_{\text{intensive margin}}}\right). \tag{36}$$

Here, consumption involve an *intensive margin*. However, as $\theta+\gamma\to\infty$, $c\left(j,t\right)\to s$ for any $j\in\left(0,n\left(t\right)\right)$; this matches the hierarchical demand with strong tilt toward lower-index varieties (they approach the satiation level s), while $c\left(j,t\right)$ falls sharply as j approaches $n(t)$.

Within-period aggregation. As in the baseline model, the expenditure function is $E\left(t\right)=\int_{0}^{n(t)}c\left(j,t\right)p\left(j,t\right)dj=R\left(t\right)\int_{0}^{n(t)}c\left(j,t\right)\hat{\varphi}\left(j|J\left(t\right)\right)dj\equiv R\left(t\right)\mathbf{e}\left(n\left(t\right);J\left(t\right)\right)$. Substituting with (36), we find the closed-form expression of the cost function for $\mathbf{e}\left(n\left(t\right);J\left(t\right)\right)=\int_{0}^{n(t)}c\left(j,t\right)\cdot\hat{\varphi}\left(j|J\left(t\right)\right)dj$, where function $\mathbf{e}\left(n;J\right)$ is given by

$$\mathbf{e}\left(n;J\right)=s\left[h_{1}\left(\frac{n}{J};\frac{\varepsilon\left(\theta+1\right)}{1-\varepsilon}-1\right)-\underbrace{\frac{1+\theta}{\left(1+\theta\right)+\left(\gamma+\theta\right)}h_{2}\left(\frac{n}{J};\frac{\varepsilon\left(2\theta+\gamma+1\right)}{1-\varepsilon}-1\right)}_{\text{intensive margin effect}}\right]\frac{n^{1+\theta}}{1+\theta},$$

with two functions h_1 and h_2 given in Appendix B. When γ is large enough, the intensive margin effect diminishes. Moreover, in the special case of $s=1$ and $\theta=0$, the first term $h_{1}\left(\frac{n}{J};\frac{\varepsilon(\theta+1)}{1-\varepsilon}-1\right)\frac{n^{1+\theta}}{1+\theta}$ collapses into the previous CES expenditure form: $n-\lambda\left[n^{\frac{\varepsilon-1}{\varepsilon}}+\left(\kappa\varepsilon\right)J^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}$. Similarly, substituting with (36), we find the closed-form expression of instantaneous utility $u\left(c\left(t\right)\right)=\left(1-\gamma\right)^{-\hat{\sigma}}\frac{c(t)^{1-\hat{\sigma}}}{1-\hat{\sigma}}$, where $c\left(t\right)$ can be expressed as a function of $n\left(t\right)$ and $J\left(t\right)$, denoted by $c\left(t\right)\equiv U\left(n\left(t\right),J\left(t\right)\right)$, and

function $U(n, J)$ is given by

$$U(n,J) = \frac{s^2}{2}\left[1 - \underbrace{\frac{\gamma-1}{(\gamma+1)+(\gamma+\theta)}\frac{1}{(n/\Delta)^{\gamma-1}-1}h_3\left(\frac{n}{J};\frac{\varepsilon(2\theta+\gamma+1)}{1-\varepsilon}-1\right)}_{\text{intensive margin effect}}\right]\frac{\Delta^{1-\gamma}-n^{1-\gamma}}{\gamma-1},$$

with function h_3 given in Appendix B. When γ is large, $(n/\Delta)^{\gamma-1}$ is large and the intensive margin effect diminishes and hence $U(n,J) \to \frac{s^2}{2}\frac{\Delta^{1-\gamma}-n^{1-\gamma}}{\gamma-1}$, which is the baseline model case.

In a nutshell, when the utility function features a strong hierarchy (i.e., γ is large enough), the extended model converges to our baseline model, so both models have the same policy implications.

Obsolescence effect. Lower-quality goods might drop out as in Stokey (1988). A slight modification to the model setup would deliver this effect. The most intuitive way is to incorporate the idea that, while learning-by-doing initially reduces the cost of a good, once the technological frontier $J(t)$ has advanced far beyond a specific good j, its production cost may begin to rise again. In other words, the unit cost function $\hat{\varphi}(j|J(t))$ is U-shaped in j for a given $J(t)$ as in Stokey (1988). For example, the unit cost function in (35) is revised by multiplying an additional "obsolescence" term $O(j, J(t)) = 1 + \pi \cdot \max(0, J(t) - j)$, where the parameter π represents the obsolescence rate. An alternative approach is to assume that the utility derived from a good is relative and depends on the current technology frontier or consumption frontier—higher-indexed goods strongly substitute for lower-indexed goods. For example, we can modify the hierarchy factor in the utility function in (34) by instead using $h(j, J(t)) = j^{-\gamma}\exp(\varrho \min(0, j - J(t)))$.

5.3 Labor

In our main model, both consumption and investment goods are produced using AK technology. Now, we consider that the production functions of consumption and investment goods are different in capital intensity. Specifically, as in Acemoglu (2009, Chapter 11.3), the production function of investment goods remains AK, whereas the production of consumption goods involves labor. We show that this does not alter the conclusions of our paper qualitatively and quantitatively.

Specifically, the production technology of investment goods remains AK with productivity z as in (10). However, the production technology of the consumption goods changes from $y(j,t) = \frac{1}{\varphi(j|J(t))}k(j,t)$ in (4) to a Cobb–Douglas production function: $y(j,t) = \frac{1}{\varphi(j|J(t))}k(j,t)^{\hat{\alpha}}\ell(j,t)^{1-\hat{\alpha}}$, where $\ell(j,t)$ represents the labor input. The labor endowment of the economy is L, with an inelastic supply. All other model settings remain unchanged from those of the baseline model.

Denote the rental rate of capital by $R(t)$ and the wage rate by $W(t)$. By the property of

Cobb–Douglas production function, it follows that for all $j \in (0, n(t)]$ at any t,

$$\frac{k(j,t) R(t)}{\ell(j,t) W(t)} = \frac{\hat{\alpha}}{1-\hat{\alpha}} \Longrightarrow \frac{k(j,t)}{\ell(j,t)} = \frac{W(t)}{R(t)} \frac{\hat{\alpha}}{1-\hat{\alpha}} = \frac{K_C(t)}{L}, \tag{37}$$

where $K_C(t) = \int_0^{n(t)} k(j,t)\, dj$ and $L = \int_0^{n(t)} \ell(j,t)\, dj$ by market clearing, recalling that $K_C(t)$ denotes the aggregate supply of capital to the consumption goods sector. Moreover, in equilibrium, $y(j,t) = 1$ for all $j \leq n(t)$, which implies $\varphi(j|J(t)) = k(j,t)^{\hat{\alpha}} \ell(j,t)^{1-\hat{\alpha}}$ for all $j \in (0, n(t)]$. Taking integral over $(0, n(t)]$ on both sides and using $\frac{k(j,t)}{\ell(j,t)} = \frac{K_C(t)}{L}$ yields

$$\int_0^{n(t)} \varphi(j|J(t)) = \int_0^{n(t)} k(j,t)^{\hat{\alpha}} \ell(j,t)^{1-\hat{\alpha}} = \left(\frac{K_C(t)}{L}\right)^{\hat{\alpha}} \int_0^{n(t)} \ell(j,t)\, dj = \left(\frac{K_C(t)}{L}\right)^{\hat{\alpha}} L.$$

We still denote $\mathbf{e}(n(t); J(t)) = \int_0^{n(t)} \varphi(j|J(t))\, dj$ and hence have $\mathbf{e}(n(t); J(t)) = \left(\frac{K_C(t)}{L}\right)^{\hat{\alpha}} L$ or $K_C(t) = \left(\frac{\mathbf{e}(n(t);J(t))}{L^{1-\hat{\alpha}}}\right)^{\frac{1}{\hat{\alpha}}}$. Therefore, the law of motion for capital changes from $\dot{k}(t) = (z-\delta) k(t) - z\mathbf{e}(n(t); J(t))$ in (19) to

$$\dot{k}(t) = (z-\delta) k(t) - z \left(\frac{\mathbf{e}(n(t); J(t))}{L^{1-\hat{\alpha}}}\right)^{\frac{1}{\hat{\alpha}}}.$$

Nothing else changes in Definition 1 (alternative); in particular, CE and SP still differ in whether they take the process of $J(t)$ as exogenously given. Clearly, the conclusion of the model does not change qualitatively. Moreover, we recalibrate this revised model with labor and find the welfare gain from the two-sided policy is 2.041%, higher than the gain in the calibrated model without labor (see Appendix B for details).

5.4 Openness

So far, our model is a closed economy model and does not consider international trade. As discussed in Murphy, Shleifer and Vishny (1989) (see the entire Section II entitled "The Importance of Domestic Markets" in their paper), there is considerable empirical evidence pointing to the overwhelming importance of the domestic market in the growth of domestic industry. The underlying reasons may be differences in tastes across countries, transportation costs and trade tariffs, and so on. Recent work by Beerli et al. (2020) documents the evidence that even for an export-oriented economy like China, the size of the domestic market is a key driver of technological progress. Doing well/improving product quality in the domestic market (e.g., through learning-by-doing) is usually the first step to entering the international market. Theoretically, Matsuyama (2019) shows that globalization amplifies, rather than reduces, the power of endogenous domestic demand composition differences as a driver of structural change. In this section, we conduct a simple extension to our

model to allow for international trade, and show that our model mechanisms can be strengthened, rather than weakened, under plausible conditions.

Consider two countries: home and rest-of-the-world (foreign). We assume that home has a comparative advantage in producing final consumption goods and a comparative disadvantage in producing capital goods, resulting in the equilibrium trade flow of home importing capital goods and exporting final goods. Within the final goods sector, only the varieties produced at the reduced cost $1-\lambda$ (i.e., $j \in (0, J(t)]$) can penetrate the international market. For each j, the foreign demand is assumed to follow $q(j,t) = \xi / p(j,t)$, where ξ is a parameter capturing the size of the foreign market. The import of capital by home is constrained by its foreign reserves, which are dependent on its export volume. The price of capital in the international market is denoted as $p_K^*(t)$, which is lower than 1 (the home price of capital). Then, the net profit of trade for the home country at time t is given by

$$\frac{\int_0^{J(t)} q(j,t)\, p(j,t)\, dj}{p_K^*(t)} - \int_0^{J(t)} q(j,t)\, \varphi(j|J(t))\, dj = \left(\frac{1}{p_K^*(t)} - 1\right) \xi J(t),$$

where the first term represents the amount of capital that the home country can purchase in the international market by using the foreign reserves earned from exports, and $p(j,t) = \varphi(j|J(t))$ represents the unit cost of producing final goods.[30] For simplicity, we assume that $p_K^*(t)$ is constant over time. Denoting $\eta^* \equiv \frac{1}{p_K^*(t)} - 1$, the net profit for the home country is

$$\Pi(t) = \xi \eta^* J(t). \tag{38}$$

It might be reasonable to assume that the home country can only export those goods that have been produced and consumed domestically. In this case, the net profit becomes

$$\Pi(t) = \xi \eta^* \min(J(t), n(t)). \tag{39}$$

With the introduction of international trade, the only change in the equilibrium characterization is the revision of the state equation in (19), which now becomes

$$\dot{k}(t) = (z - \delta) k(t) - z \hat{\mathbf{e}}(n(t); J(t)),$$

where $\hat{\mathbf{e}}(n(t); J(t)) = \mathbf{e}(n(t); J(t)) - \Pi(t)/z$.[31] Note that, under (39), the expenditure function

[30] We assume competitive home exporters and also implicitly assume that they cannot directly tap the international capital market due to frictions such as global sourcing (see, e.g., Antràs and Helpman (2008)).

[31] The term $\Pi(t)$ can be interpreted more broadly. For example, it can also be interpreted as the profit on processing trade or the net capital return on savings aboard under unbalanced trade, in which case export is effectively equivalent to more domestic capital accumulation in our closed-economy model.

$\hat{\mathbf{e}}(n(t); J(t))$ can be written in exactly the same form as in (25) but with the revised parameter $\lambda^* = \lambda + \frac{\xi\eta^*}{z}$.

It is easy to see that with international trade, the competitive equilibrium and the social planner's solution studied in Section 2 remain qualitatively unchanged, under either (38) or (39). However, the presence of international trade amplifies the loop between $\{J(t)\}_{t=0}^{\infty}$ and $\{N(t)\}_{t=0}^{\infty}$. Specifically, the net capital inflow from international trade accelerates domestic consumption and industrial upgrading, which leads to a wider range of consumption goods becoming cheaper and exportable, resulting in even more capital inflow. The amplification is reflected by the effect that a small increase in ξ (the size of foreign market) can lead to a large increase in $(J(t), N(t))$ at a given time point t. This means that not only the welfare is improved but also the steady-state growth rate can be upgraded to a higher level based on the result in Section 3. It is also worth noting that under (39), international trade results in a higher effective λ, in which case our result of two-phase intervention in Proposition 2 becomes stronger.

6 Conclusion

This paper develops a dynamic model with hierarchy-of-needs preference and learning-by-doing production. In a growth setting, the approach we follow allows for analytical transparency on the interactions between the supply- and the demand-side forces. We believe that our paper adds to the understanding of the mechanisms of (optimal) development.

A large extant literature has pointed out that production externalities and spillovers (or simply learning-by-doing) are a fundamental force behind economic growth. We emphasize that for a catching-up economy, production externalities and spillovers largely occur in the consumer goods sector as in Stokey (1988) and Matsuyama (2002), at least at certain stages of development. The model helps to shed light on some macroeconomic phenomena in developing countries. In particular, the model has implications for optimal development policies. For some, a mass consumption society is viewed as the mature stage of development (e.g., Rostow (1959)). However, history tells us that successfully transitioning to this stage is rare: it simply does not arise or, when it does, it fails to be sustained. Our paper provides one perspective of inquiry.

Reference

Acemoglu, D. (1998). Why do New Technologies Complement Skills? Directed Technical Change and the Wage Inequality, Quarterly Journal of Economics, 113, 1055-1089.

Acemoglu, D. (2002). Directed Technical Change, Review of Economic Studies, 781-809.

Acemoglu, D. (2009). Introduction to Modern Economic Growth, Princeton Press.

Acemoglu, D., Aghion, P. & Zilibotti, F. (2006). Distance to Frontier, Selection, and Economic Growth, Journal of the European Economic Association, 4, 1, 37–74.

Acemoglu, D. & Linn, J. (2004). Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry, Quarterly Journal of Economics, 119 (3), 1049-1090.

Acemoglu, D., & Zilibotti, F. (2001). Productivity Differences, Quarterly Journal of Economics, 116, 563-606.

Aghion P., Barrage, L., Hémous, D., & Liu, E. (2024). Transition to Green Technology along the Supply Chain, working paper.

Aghion, P., & Howitt, P. (1992). A Model of Growth Through Creative Destruction. Econometrica, 60(2), 323–351.

Akcigit, U., Wang, Y., Xu, W., & Zhu, L. (2025). Leapfrogging by Switching Lanes: Directed Technical Change in China's Transition to Electric Vehicles, working paper.

Arrow, K. J. (1962). The Economic Implications of Learning by Doing. The Review of Economic Studies, 29(3), 155–173.

Banerjee, A, & E. Duflo (2005). Growth theory through the lens of development economics. In: Handbook of Economic Growth 1, 473–552.

Barro, R. J. (2003). Determinants of economic growth in a panel of countries. Annals of economics and finance.

Barwick, P, Kwon, H.S, Li, S & Zahur, N.B (2025). Drive Down the Cost: Learning by Doing and Government Policies in the Global EV Battery Industry, NBER w33378.

Beerli, A., Weiss, F. J., Zilibotti, F., & Zweimüller, J. (2020). Demand forces of technical change evidence from the Chinese manufacturing industry, China Economic Review, 60(C).

Benhabib, J., Perla, J., & Tonetti, C. (2014). Catch-Up and Fall-Back through Innovation and Imitation. Journal of Economic Growth, 19(1), 1–35.

Bergman, N., Jaimovich, N., & Saporta-Eksten, I. (2025). The Impact of Economic Shifts: From Micro-Elasticities to Distributional Welfare Effects, working paper.

Bertola, G., Foellmi, R., & Zweimüller, J. (2014). Income Distribution in Macroeconomic Models. Princeton University Press.

Boissay, F., Collard, F., & Smets, F. (2016). Booms and Banking Crises, Journal of Political Economy, 124, 2, 489-538.

Boppart, T. (2014). Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences, Econometrica, 82(6), 2167–2196.

Boppart, T. & Krusell, P. (2020). Labor Supply in the Past, Present, and Future: A Balanced-Growth Perspective, Journal of Political Economy, 128(1), 118-157.

Boppart, T. & Weiss, F. J. (2013). Non-homothetic preferences and industry directed technical change, working paper.

Brunnermeier, M., Gourinchas, P., & Itskhoki, O. (2018). Consumption-led Growth.

Buera, F. J., Hopenhayn, H., Shin, Y., & Trachter, N. (2021). Big Push in Distorted Economies. NBER Working Paper 28561.

Buera, F. J., & Kaboski, J. P. (2012). The Rise of the Service Economy. American Economic Review, 102(6), 2540–2569.

Buera, F. J., & Shin, Y. (2013). Financial Frictions and the Persistence of History: A Quantitative Exploration, Journal of Political Economy, 121, 2, 221-272.

Cai, J. & Li, N. (2019). Growth Through Inter-sectoral Knowledge Linkages, Review of Economic Studies, 86 (5), 1827-1866.

Comin, D., Lashkari, D., & Mestieri, M. (2021). Structural change with long-run income and price effects. Econometrica, 89(1): 311–374.

Dávila, E. & Korinek, A. (2018). Pecuniary Externalities in Economies with Financial Frictions, Review of Economic Studies, 85, 1, 352–395.

Dávila, E. & Schaab, A. (2023). Welfare Assessments with Heterogeneous Individuals, working paper.

Dávila, E. & Walther A. (2022). Corrective Regulation with Imperfect Instruments, working paper.

Foellmi, R., & Zweimüller, J. (2006). Income Distribution and Demand-Induced Innovations, Review of Economic Studies, 73, 4, 941–960.

Foellmi, R., & Zweimüller, J. (2008). Structural Change, Engel's Consumption Cycles and Kaldor's Facts of Economic Growth. Journal of Monetary Economics, 55(7), 1317–1328.

Goldberg, P.K. & Reed, T. (2023). Presidential Address: Demand-Side Constraints in Development. The Role of Market Size, Trade, and (In)Equality. Econometrica, 91,1915-50.

Greenwood, J., Hercowitz, Z., & Krusell, P. (2000). The role of investment-specific technological change in the business cycle. European Economic Review, 44(1):91–115.

Greenwood, J. & Uysal, G. (2005). New goods and the transition to a new economy. Journal of Economic Growth, 10(2):99–134.

Griliches, Z. (1957). Hybrid Corn: An Exploration in the Economics of Technological Change. Econometrica, 25(4), 501–522.

Grossman, G.M. & Helpman, E. (1991). Innovation and Growth in the Global Economy. MIT Press, Cambridge, MA.

Halter, D., Oechslin, M., & Zweimüller, J. (2014). Inequality and growth: the neglected time dimension. Journal of Economic Growth, 19(1):81–104.

Itskhoki, O. & Moll, B. (2019). Optimal Development Policies With Financial Frictions. Econometrica, 87, 1, 139-173.

Jaimovich, N., & Rebelo, S. (2017). Nonlinear Effects of Taxation on Growth. Journal of Political Economy, 125(1), 306–340.

Kamien, M.I. & Schwartz, N.L. (2012). Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management.

Kongsamut, P., Rebelo, S., & Xie, D. (2001). Beyond Balanced Growth. The Review of Economic Studies, 68(4), 869–882.

Krugman, P. (1979). A Model of Innovation, Technology Transfer, and the World Distribution of Income, Journal of Political Economy, 87, 2, 253-266.

Krugman, P. (1991). History Versus Expectations, Quarterly Journal of Economics, 106, 2, 651-667.

Lamoreaux, N., Raff, D., & Temin, P. (1998). Learning by Doing in Markets, Firms, and Countries, University of Chicago Press.

Liu, E. (2019). Industrial Policies in Production Networks. Quarterly Journal of Economics, 134(4), 1883–1948.

Liu, E. & Ma, S. (2022). Innovation Networks and R&D Allocation, working paper.

Lorenzoni, G. (2008). Inefficient Credit Booms, Review of Economic Studies, 75, 3, 809–833.

Lucas, R. (1993). Making a Miracle. Econometrica, 61, 2, 251-272.

Lucking, B., Bloom, N. & Van Reenen, J. (2019). Have R&D Spillovers Declined in the 21st Century? Fiscal Studies, 40 (4), 561-590.

Matsuyama, K. (1992a). Agricultural Productivity, Comparative Advantage, and Economic Growth, Journal of Economic Theory 58, 317-334.

Matsuyama, K. (1992b). The market size, entrepreneurship, and the big push, Journal of the Japanese and International Economies, 6(4), 347-364.

Matsuyama, K. (1995). Complementarities and Cumulative Processes in Models of Monopolistic Competition. Journal of Economic Literature, 33(2), 701–729.

Matsuyama, K. (2002). The Rise of Mass Consumption Societies. Journal of Political Economy, 110(5), 1035–1070.

Matsuyama, K. (2009). Structural Change in an Interdependent World: A Global View of Manufacturing Decline, Journal of the European Economic Association, 7, 478-486.

Matsuyama, K. (2019). Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade. Econometrica, 87(2), 497–528.

Murphy, K. M., Shleifer, A. & Vishny, R. W. (1989). Industrialization and the Big Push. Journal of Political Economy, 97(5), 1003–1026.

Oberfield, E. (2023). Inequality and Measured Growth. working paper.

Perla, J., & Tonetti, C. (2014). Equilibrium Imitation and Growth. Journal of Political Economy, 122(1), 52–76.

Rebelo, S. (1991). Long-Run Policy Analysis and Long-Run Growth. Journal of Political Economy, 99(3), 500–521.

Rodrik, D. (2004). Industrial Policy for the Twenty-First Century. Cambridge, MA: Harvard University.

Rodrik, D. (2023). A Comment on "Presidential Address: Demand-Side Constraints in Development: The Role of Market Size, Trade, and (In)Equality".

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. Journal of Political Economy, 94(5), 1002-1037.

Romer, P. M. (1990). Endogenous Technological Change. Journal of Political Economy, 98(5), 71-102.

Rostow, W. (1959). The Stages of Economic Growth. Economic History Review, 12(1).

Schmookler, J. (1966). Invention and Economic Growth, Harvard University Press.

Stokey, N. L. (1988). Learning by Doing and the Introduction of New Goods. Journal of Political Economy, 96(4), 701–717.

Young, A. (1991). Learning by Doing and the Dynamic Effects of International Trade, Quarterly Journal of Economics, 106, 2, 369-405.

Young, A. (1993). Invention and Bounded Learning by Doing, Journal of Political Economy, 101(3), 443-472.

Zilibotti, F. (2017). Growing and Slowing Down Like China, Journal of the European Economic Association,15(5), 943-988.

Zweimüller, J. (2000). Schumpeterian Entrepreneurs Meet Engel's Law: The Impact of Inequality on Innovation-Driven Growth, Journal of Economic Growth, 5, 2, 185-206.

A Proofs

Proof of Proposition 1: We solve competitive equilibrium in several steps.

Step 1: $J(t)$ must be a continuous process without jumps.

This is straightforwardly obtained. In fact, based on (26), $J(t)$ is a continuous process even if $n(t)$ is not a continuous process.

Step 2: $n(t)$ must be a continuous process without jumps, and there exist two thresholds τ_1 and τ_2 such that $n(t)\begin{cases} < J(t) & \text{for } t \in [0, \tau_1) \\ = J(t) & \text{for } t \in [\tau_1, \tau_2] \\ > J(t) & \text{for } t \in (\tau_2, +\infty) \end{cases}$.

In CE, individual households take the J process as exogenously given and treat the entire process $\{J(s)\}_{s=t}^{\infty}$ rather than $J(t)$ as a state variable at time t. When K_0 is low enough relative to J_0, we can verify that there are three stages (regimes) for $n(t)$. Following Kamien and Schwartz (2012) to study a "switching" problem, the *current value* Hamiltonian can be formulated as

$\mathcal{H}_3 = \frac{1}{1-\sigma} n^{1-\sigma} + m\left[(z-\delta)k - zn + z\lambda J\right]$, for $t \in (\tau_2, +\infty)$

$\mathcal{H}_2 = \frac{1}{1-\sigma} n^{1-\sigma} + m\left[(z-\delta)k - z(1-\lambda)n\right]$, for $t \in (\tau_1, \tau_2)$

$\mathcal{H}_1 = \frac{1}{1-\sigma} n^{1-\sigma} + m\left[(z-\delta)k - z(1-\lambda)n\right]$, for $t \in [0, \tau_1)$.

The optimal conditions within each stage imply

- stage 3: $t \in (\tau_2, +\infty)$

  $n^{-\sigma} - mz = 0$; $\frac{\dot{m}}{m} = -(z - \delta - \rho)$

- stage 2: $t \in (\tau_1, \tau_2)$

  $n = J$ and $n^{-\sigma} - mz(1-\lambda) \geq 0$ (corner solution); $\frac{\dot{m}}{m} = -(z - \delta - \rho)$

- stage 1: $t \in [0, \tau_1)$

  $n^{-\sigma} - mz(1-\lambda) = 0$; $\frac{\dot{m}}{m} = -(z - \delta - \rho)$.

We find the optimal conditions at "switching" points, τ_1 and τ_2. The first-order condition of V_2 with respect to τ_2 evaluated at $t = \tau_2^-$ is

$$\left.\frac{\partial V_2(t, \tau_2)}{\partial \tau_2}\right|_{t=\tau_2^-} = \frac{1}{1-\sigma} J(\tau_2)^{1-\sigma} + m^+(\tau_2) \left[\underbrace{(z-\delta)k - z(1-\lambda)J}_{\text{state equation before } \tau_2}\right] - \rho V_3(\tau_2)$$

$$= \frac{1}{1-\sigma}\left(J(\tau_2)^{1-\sigma} - n^+(\tau_2)^{1-\sigma}\right) - m^+(\tau_2) z\left(J(\tau_2) - n^+(\tau_2)\right), \text{ (by using } \mathcal{H}_3(\tau_2) = \rho V_3(\tau_2)\text{)}$$

which is equal to 0 if and only if $n^{+}(\tau_2) = J(\tau_2)$, by applying the mean value theorem with the first-order condition $n^{-\sigma} - mz = 0$ at τ_2^{+}. Hence, the control variable n is continuous at τ_2. By using $\mathcal{H}_2(\tau_2) = \mathcal{H}_3(\tau_2)$, we also obtain that m is continuous at τ_2, that is, $m^{-}(\tau_2) = m^{+}(\tau_2)$. Similarly, it is easy to show that n and m are continuous at τ_1. Then, by $J(\tau_2)^{-\sigma} - m(\tau_2) z = 0$, $J(\tau_1)^{-\sigma} - m(\tau_1) z(1-\lambda) = 0$, and $\frac{\dot{m}}{m} = -(z - \delta - \rho)$, we find

$$\tau_2 - \tau_1 = \frac{-\log(1-\lambda)}{(z-\delta-\rho) - \sigma b}. \tag{A.1}$$

Step 3: We find out the joint processes $\{N(t), J(t)\}$ for $t \geq \tau_2$.

From Step 2, we know that at stage 3, the growth rate of consumption is constant. So, if we know $n(t)$ at some $t \in [\tau_2, +\infty)$, we can express the consumption path afterwards as $n(s) = n(t) e^{\frac{(z-\delta)-\rho}{\sigma}(s-t)}$. This in turn gives the process of $J(t)$ (by (26)), that is, $J(s) = \left(J(t) - \frac{\phi n(t)^{\alpha}}{\frac{(z-\delta)-\rho}{\sigma}\alpha - b}\right) e^{b(s-t)} + \frac{\phi n(t)^{\alpha}}{\frac{(z-\delta)-\rho}{\sigma}\alpha - b} e^{\left[\frac{(z-\delta)-\rho}{\sigma}\alpha\right](s-t)}$. Then we can define the lifetime wealth ("endowment") created by the technology process starting from t, namely,

$$\Lambda(t) = z\lambda \int_t^{\infty} \exp\left(-(s-t)(z-\delta)\right) J(s)\, ds = z\lambda \left[\frac{\left(J(t) - \frac{\phi n(t)^{\alpha}}{\frac{(z-\delta)-\rho}{\sigma}\alpha - b}\right)}{(z-\delta) - b} + \frac{\frac{\phi n(t)^{\alpha}}{\frac{(z-\delta)-\rho}{\sigma}\alpha - b}}{(z-\delta) - \frac{(z-\delta)-\rho}{\sigma}\alpha} \right]. \tag{A.2}$$

Based on the budget constraint in Program (19), we can write the lifetime budget constraint of a household starting from time $t \in [\tau_2, +\infty)$ as $k(t) + \Lambda(t) = \int_t^{+\infty} \exp\left(-(s-t)(z-\delta)\right) \cdot z n(s)\, ds$, which is essentially a transversality condition. Plugging $n(s) = n(t) e^{\frac{(z-\delta)-\rho}{\sigma}(s-t)}$ into the lifetime budget constraint, we have

$$n(t) = (k(t) + \Lambda(t)) \left[(z-\delta) - \frac{(z-\delta)-\rho}{\sigma} \right] \frac{1}{z}. \tag{A.3}$$

The two equations (A.2) and (A.3), forming a fixed-point problem, jointly solve the two unknowns $\Lambda(t)$ and $n(t)$. Thus, both $\Lambda(t)$ and $n(t)$ are expressed as a function of the state variables $k(t)$ and $J(t)$.

Specifically, $\Lambda(K, J)$ is the solution to the implicit function

$$\Lambda = z\lambda \left[\frac{J}{z-\delta} + \frac{\phi \left(\left[(z-\delta) - \frac{(z-\delta)-\rho}{\sigma} \right] \frac{1}{z} \right)^{\alpha}}{(z-\delta)\left[(z-\delta) - \frac{(z-\delta)-\rho}{\sigma}\alpha \right]} (K+\Lambda)^{\alpha} \right]. \tag{A.4}$$

Furthermore, we can calculate the value function at $t \in [\tau_2, +\infty)$ as

$$V_3 = V_3\left(k, \Lambda\left(K, J\right)\right) = \frac{1}{1-\sigma}\left[\left(z-\delta\right) - \frac{\left(z-\delta\right)-\rho}{\sigma}\right]^{-\sigma} z^{\sigma-1}\left[k + \Lambda\left(K, J\right)\right]^{1-\sigma}, \tag{A.5}$$

where k denotes individual agents' state variable and K denotes the aggregate state variable. This also implies that $m_{\text{CE}} = \frac{\partial V_3}{\partial k}$ and $m_{\text{SP}} = \frac{\partial V_3}{\partial k} + \Lambda_K \frac{\partial V_3}{\partial \Lambda}$, where m is the costate variable (i.e., the marginal valuation of capital).

Step 4: We find out the properties of the two endogenous thresholds τ_1 and τ_2.

The results (A.2) and (A.3) derived in Step 3 apply to any $t \in [\tau_2, +\infty)$. In particular, they are true at the boundary $t = \tau_2$. As we also know $n\left(\tau_2^-\right) = J\left(\tau_2^-\right)$, we hence obtain a relationship between $k(t)$ and $J(t)$ at $t = \tau_2$. Specifically, the condition $n\left(\tau_2^-\right) = J\left(\tau_2^-\right)$ together with (A.2) and (A.3) imply that at $t = \tau_2$ the relationship $k\left(\tau_2\right) = K^{**}\left(J\left(\tau_2\right)\right)$ must hold, where function $K^{**}\left(\cdot\right)$ is *uniquely* defined as

$$K^{**}\left(J\right) = \frac{zJ}{\left(z-\delta\right) - \frac{\left(z-\delta\right)-\rho}{\sigma}} - z\lambda \cdot \left(\frac{J}{\left(z-\delta\right)-b} + \frac{\phi J^{\alpha}}{\left[\left(z-\delta\right)-b\right]\left[\left(z-\delta\right) - \frac{\left(z-\delta\right)-\rho}{\sigma}\alpha\right]}\right). \tag{A.6}$$

For given $k(t)$ and $J(t)$ at some $t \in [\tau_1, \tau_2)$ in the second stage, we can find out the evolution of $k(t)$ and $J(t)$ based on the results in Step 1. In fact, $J(s) = J(t)\exp\left[b\left(s-t\right)\right]$ and $k(s) = \left(k(t) - \frac{z(1-\lambda)J(t)}{\left(z-\delta\right)-b}\left[1 - \exp\left[-\left(\left(z-\delta\right)-b\right)\left(s-t\right)\right]\right]\right)\exp\left[\left(z-\delta\right)\left(s-t\right)\right]$ for $s \in [t, \tau_2]$. Substituting $s = \tau_2$ and using the result $k\left(\tau_2\right) = K^{**}\left(J\left(\tau_2\right)\right)$ yields

$$\begin{aligned} &\left(k(t) - \frac{z\left(1-\lambda\right)J\left(t\right)}{\left(z-\delta\right)-b}\left[1 - \exp\left[-\left(\left(z-\delta\right)-b\right)\left(\tau_2-t\right)\right]\right]\right)\exp\left[\left(z-\delta\right)\left(\tau_2-t\right)\right] \\ = \;& \frac{J(t)\exp\left(b\left(\tau_2-t\right)\right)}{\left(z-\delta\right) - \frac{\left(z-\delta\right)-\rho}{\sigma}} z - z\lambda\left(\frac{J(t)\exp\left[b\left(\tau_2-t\right)\right]}{\left(z-\delta\right)-b} + \frac{\phi\left[J(t)\exp\left(b\left(\tau_2-t\right)\right)\right]^{\alpha}}{\left[\left(z-\delta\right)-b\right]\left[\left(z-\delta\right) - \frac{\left(z-\delta\right)-\rho}{\sigma}\alpha\right]}\right). \end{aligned} \tag{A.7}$$

Setting $t = \tau_1$ in the above equation and using the result $\tau_2 - \tau_1 = \frac{-\log(1-\lambda)}{\left(z-\delta-\rho\right)-\sigma b}$ in (A.1), we obtain a relationship between $k(t)$ and $J(t)$ at $t = \tau_1$, that is, at $t = \tau_1$ the condition $k\left(\tau_1\right) = K^{*}\left(J\left(\tau_1\right)\right)$ must be satisfied, where function $K^{*}\left(\cdot\right)$ is given by

$$K^{*}\left(J\right) = \frac{\frac{Jz\exp\left(b\frac{-\log(1-\lambda)}{\left(z-\delta-\rho\right)-\sigma b}\right)}{\left(z-\delta\right)-\frac{\left(z-\delta\right)-\rho}{\sigma}} - \frac{Jz}{\left(z-\delta\right)-b}\left(1-\lambda\right)^{\frac{-b}{\left(z-\delta-\rho\right)-\sigma b}} - \frac{z\lambda\cdot\phi\left[J\exp\left(b\frac{-\log(1-\lambda)}{\left(z-\delta-\rho\right)-\sigma b}\right)\right]^{\alpha}}{\left[\left(z-\delta\right)-b\right]\left[\left(z-\delta\right)-\frac{\left(z-\delta\right)-\rho}{\sigma}\alpha\right]}}{\left(1-\lambda\right)^{\frac{-\left(z-\delta\right)}{\left(z-\delta-\rho\right)-\sigma b}}} + \frac{z\left(1-\lambda\right)J}{\left(z-\delta\right)-b}. \tag{A.8}$$

Step 5: We find out the initial consumption $n\left(0\right)$ and τ_1.

For the given initial state $(k(0), J(0))$, if $n(0)$ is known, we can find the processes of $k(t)$ and $J(t)$ at $t \in [0, \tau_1]$. That is, $k(t) = \left[k(0) - \frac{n(0)z(1-\lambda)}{(z-\delta)-\frac{z-\delta-\rho}{\sigma}}\left(1 - e^{-\left[(z-\delta)-\frac{z-\delta-\rho}{\sigma}\right]t}\right)\right] e^{(z-\delta)t}$ and $J(t) = J(0)e^{bt}$. Letting $t = \tau_1$ and using the relationship $k(\tau_1) = K^*(J(\tau_1))$ yields the equation in the result (2) of Proposition 1. This equation together with $J(\tau_1) = J(0)\exp(b\tau_1)$ jointly solve the two unknowns $n(0)$ and τ_1.

Proof of Corollary 1: By $\frac{\dot{J}(t)}{J(t)} = b + \phi\frac{N(t)^\alpha}{J(t)}$, it follows $\lim_{t\to\infty}\frac{J(t)}{N(t)^\alpha} = \frac{\phi}{\alpha\frac{(z-\delta)-\rho}{\sigma}-b}$. By $\frac{\dot{k}(t)}{k(t)} = (z-\delta) - z\frac{n(t)}{k(t)} + z\lambda\frac{J(t)}{k(t)}$, it follows $\lim_{t\to\infty}\frac{N(t)}{K(t)} = \frac{(z-\delta)-\frac{(z-\delta)-\rho}{\sigma}}{z}$. We also need to have $g_J(\tau_2) \leq g_N(\tau_2)$, which requires the parameter condition $b + \phi J_0^{\alpha-1} \leq \frac{(z-\delta)-\rho}{\sigma}$.

Proof of Proposition 2: i) We solve the social planner problem. We consider the case where K_0 is low enough relative to J_0. In this case, when increasing n from J to above J, the planner faces a jump in the marginal cost of consumption but also a jump in the impact of n on J in (26). For parameter values such that the former dominates the latter (we will give the condition later), the equilibrium process is continuous with three stages.

There exist two thresholds τ_1 and τ_2 such that $n(t)\begin{cases} < J(t) & \text{for } t \in [0, \tau_1) \\ = J(t) & \text{for } t \in (\tau_1, \tau_2) \\ > J(t) & \text{for } t \in (\tau_2, +\infty) \end{cases}$. The current value Hamiltonian can be formulated as

$\mathcal{H}_3 = \frac{1}{1-\sigma}n^{1-\sigma} + m[(z-\delta)k - zn + z\lambda J] + \mu(\phi n^\alpha)$, for $t \in [\tau_2, +\infty)$

$\mathcal{H}_2 = \frac{1}{1-\sigma}n^{1-\sigma} + m[(z-\delta)k - z(1-\lambda)n]$, for $t \in (\tau_1, \tau_2)$

$\mathcal{H}_1 = \frac{1}{1-\sigma}n^{1-\sigma} + m[(z-\delta)k - z(1-\lambda)n]$, for $t \in [0, \tau_1]$.

The optimal conditions within each stage imply

- stage 3: $t \in (\tau_2, +\infty)$

  $n^{-\sigma} - mz + \mu\phi\alpha n^{\alpha-1} = 0$; $\frac{\dot{m}}{m} = -(z-\delta-\rho)$; $-\dot{\mu} = mz\lambda + \mu(-\rho)$

- stage 2: $t \in (\tau_1, \tau_2)$

  $n = J$ and $n^{-\sigma} - mz(1-\lambda) \geq 0$ (corner solution); $\frac{\dot{m}}{m} = -(z-\delta-\rho)$

- stage 1: $t \in [0, \tau_1)$

  $n^{-\sigma} - mz(1-\lambda) = 0$; $\frac{\dot{m}}{m} = -(z-\delta-\rho)$

We next examine the conditions at the switching points, τ_1 and τ_2. We prove that the control (policy) variable n is continuous at time τ_2 whereas the costate variable m is not continuous at τ_2.

Note that $J(t)$ is continuous at $t=\tau_2$ (even if $n(t)$ is not a continuous at $t=\tau_2$). The first-order condition of V_2 with respect to τ_2 evaluated at $t=\tau_2^-$ is

$$\left.\frac{\partial V_2(t,\tau_2)}{\partial \tau_2}\right|_{t=\tau_2^-} = \frac{1}{1-\sigma} J(\tau_2)^{1-\sigma} + m^+(\tau_2)\left[(z-\delta)k - z(1-\lambda)J\right] - \rho V_3(\tau_2).$$

Substituting $\rho V_3(\tau_2) = \mathcal{H}_3(\tau_2)$ (by HJB) yields

$$\left.\frac{\partial V_2(t,\tau_2)}{\partial \tau_2}\right|_{t=\tau_2^-} = \left[\begin{array}{l} \frac{1}{1-\sigma}\left(J(\tau_2)^{1-\sigma} - n^+(\tau_2)^{1-\sigma}\right) - m^+(\tau_2) z\left(J(\tau_2) - n^+(\tau_2)\right) \\ +\mu^+(\tau_2)\left(\phi J(\tau_2)^{\alpha} - \phi n^+(\tau_2)^{\alpha}\right) \end{array}\right] - \mu^+(\tau_2)\phi J(\tau_2)^{\alpha} < 0$$

where the last inequality follows by applying the mean value theorem with the first-order condition $n^{-\sigma} - mz + \mu\phi\alpha n^{\alpha-1} = 0$ at $t=\tau_2^+$. This means that the planner wants to choose a τ_2 that is as low as possible until the constraint $n^+(\tau_2) \geq J(\tau_2)$ becomes binding. So the control (policy) variable n is continuous at time τ_2. Then, by using $\mathcal{H}_2(\tau_2) = \mathcal{H}_3(\tau_2)$ and $n^+(\tau_2) = n^-(\tau_2)$, we obtain

$$m^+(\tau_2) - m^-(\tau_2) = \frac{-\mu^+(\tau_2)\phi n(\tau_2)^{\alpha}}{(z-\delta)k(\tau_2) - z n(\tau_2) + z\lambda J(\tau_2)}, \tag{A.9}$$

that is, the costate variable m is not continuous at τ_2; the jump reflects the externality. Moreover, it is easy to show that n and m are continuous at time τ_1.

The transversality conditions are $\lim_{t\to\infty} \exp(-\rho t) k(t) m(t) = 0$ and $\lim_{t\to\infty} \exp(-\rho t) J(t) \mu(t) = 0$. We then prove that in the third stage, for any $t \in [\tau_2, +\infty)$, the relationship $\frac{\mu(t)}{m(t)} = \frac{z\lambda}{z-\delta}$ holds. Fix a time point $s \in [\tau_2, +\infty)$ and consider the process afterwards at $t \geq s$. We have $\mu(t) = \left[\mu(s) + \frac{m(s) z\lambda}{(-\rho)-(z-\delta-\rho)}\left(1 - \exp\left(\left[(-\rho) - (z-\delta-\rho)\right](t-s)\right)\right)\right]\exp(-(-\rho)(t-s))$, by $\frac{\dot{m}}{m} = -(z-\delta-\rho)$ and $\dot{\mu} = -mz\lambda - \mu(-\rho)$. Substituting $\mu(t)$ into $\lim_{t\to\infty} \exp(-\rho(t-s)) J(t) \mu(t) = 0$ yields

$$\lim_{t\to\infty} J(t) \underbrace{\left(\mu(s) - \frac{m(s) z\lambda}{z-\delta}\right)}_{\text{constant}} + \underbrace{J(t) \frac{m(s) z\lambda}{z-\delta} \exp\left(-(z-\delta)(t-s)\right)}_{\to 0} = 0,$$

which implies $\frac{\mu(s)}{m(s)} = \frac{z\lambda}{z-\delta}$. Then, the first-order condition $n^{-\sigma} - mz + \mu\phi\alpha n^{\alpha-1} = 0$ is replaced by $n^{-\sigma} = mz\left(1 - \frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}\right)$, and the costate variable μ is eliminated.

To summarize the solution for the third stage, for a given state variable $(k(t), J(t))$ at t in the third stage, we can find the corresponding $(m(t), n(t))$ at t. The equation system is summarized

here.

Social planner	Competitive equilibrium
$\frac{\dot{m}_{\mathrm{SP}}}{m_{\mathrm{SP}}} = -(z-\delta-\rho)$	$\frac{\dot{m}_{\mathrm{CE}}}{m_{\mathrm{CE}}} = -(z-\delta-\rho)$
$n^{-\sigma} = m_{\mathrm{SP}} z\left(1-\frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}\right)$	$n^{-\sigma} = m_{\mathrm{CE}} z$
$\dot{k} = (z-\delta)k - zn + z\lambda J$	$\dot{k} = (z-\delta)k - zn + z\lambda J$
$\dot{J} = bJ + \phi n^{\alpha}$	$\dot{J} = bJ + \phi n^{\alpha}$
$\left\{\begin{array}{l} K(t) - z\int_t^{+\infty} \exp\left(-(z-\delta)(s-t)\right)\cdot n(s)\,ds \\ +z\lambda\int_t^{+\infty} \exp\left(-(z-\delta)(s-t)\right)\cdot J(s)\,ds = 0 \end{array}\right\}$	$\left\{\begin{array}{l} K(t) - z\int_t^{+\infty} \exp\left(-(z-\delta)(s-t)\right)\cdot n(s)\,ds \\ +z\lambda\int_t^{+\infty} \exp\left(-(z-\delta)(s-t)\right)\cdot J(s)\,ds = 0 \end{array}\right\}$

We can see the only difference between SP and CE is the first-order condition with respect to n.

Then using the condition $n(\tau_2) = J(\tau_2)$, we find a relationship between $k(t)$ and $J(t)$ at $t = \tau_2$; that is, we find a function $K^{**}(\cdot)$ which satisfies $k(\tau_2) = K^{**}(J(\tau_2))$. After finding function $K^{**}(\cdot)$, we go back to stages 1 and 2 to find five unknowns $(n(0), \tau_1, \tau_2, k(\tau_1), k(\tau_2))$ for a given $(K(0), J(0))$ by jointly solving the following five equations:

1) $k(\tau_2) = K^{**}(J(0))$

2) $\tau_2 - \tau_1 = \frac{1}{z-\delta-\rho}\log\frac{m(\tau_1)}{m^-(\tau_2)}$

where $n(\tau_1)^{-\sigma} - m(\tau_1)z(1-\lambda) = 0$, $m^-(\tau_2)$ is given by (A.9), and $m^+(\tau_2)$ satisfies $n^{-\sigma} - mz + \mu\phi\alpha n^{\alpha-1} = 0$,

3) $k(\tau_2) = \left(k(\tau_1) - \frac{z(1-\lambda)n(\tau_1)}{z-\delta}\left[1-\exp\left[-(z-\delta)(\tau_2-\tau_1)\right]\right]\right)\exp\left[(z-\delta)(\tau_2-\tau_1)\right]$,
where $n(\tau_1) = J(\tau_1) = J(0)$,

4) $n(\tau_1) = n(0)\cdot\exp\left(\frac{z-\delta-\rho}{\sigma}\tau_1\right)$, where $n(\tau_1) = J(\tau_1) = J(0)$,

5) $k(\tau_1) = \left(k(0) - n(0)\frac{z(1-\lambda)}{(z-\delta)-\frac{z-\delta-\rho}{\sigma}}\left[1-\exp\left[-\left((z-\delta)-\frac{z-\delta-\rho}{\sigma}\right)\tau_1\right]\right]\right)\exp\left[(z-\delta)\tau_1\right]$.

It is important to note that in solving the problem of the first two stages, the only difference between CE and SP is the condition to determine equilibrium $\tau_2 - \tau_1$. For CE, $\tau_2 - \tau_1$ is given by (A.1). For SP, $m^-(\tau_2) = m^+(\tau_2) + \frac{\mu^+(\tau_2)\phi n(\tau_2)^{\alpha}}{\dot{k}(\tau_2)} = \left(1+\frac{\phi n(\tau_2)^{\alpha}}{\dot{k}(\tau_2)}\frac{z\lambda}{z-\delta}\right)m^+(\tau_2) = \left(1+\frac{\phi n(\tau_2)^{\alpha}}{\dot{k}(\tau_2)}\frac{z\lambda}{z-\delta}\right)\frac{n(\tau_2)^{-\sigma}/z}{1-\frac{\lambda}{z-\delta}\phi\alpha n(\tau_2)^{\alpha-1}} = \left(1+\frac{\phi n(\tau_2)^{\alpha}}{\dot{k}(\tau_2)}\frac{z\lambda}{z-\delta}\right)\frac{m_{\mathrm{CE}}(\tau_2)}{1-\frac{\lambda}{z-\delta}\phi\alpha n(\tau_2)^{\alpha-1}}$, and

$$\begin{aligned}\tau_2 - \tau_1 &= \frac{1}{z-\delta-\rho}\log\frac{m(\tau_1)}{m^-(\tau_2)} = \frac{1}{z-\delta-\rho}\log\frac{n(\tau_1)^{-\sigma}/(z(1-\lambda))}{\left(1+\frac{\phi n(\tau_2)^{\alpha}}{\dot{k}(\tau_2)}\frac{z\lambda}{z-\delta}\right)\frac{n(\tau_2)^{-\sigma}/z}{1-\frac{\lambda}{z-\delta}\phi\alpha n(\tau_2)^{\alpha-1}}} \\ &= \frac{-\log(1-\lambda)}{z-\delta-\rho} + \frac{1}{z-\delta-\rho}\log\frac{1-\frac{\lambda}{z-\delta}\phi\alpha n(\tau_2)^{\alpha-1}}{1+\frac{z\lambda}{z-\delta}\frac{\phi n(\tau_2)^{\alpha}}{\dot{k}(\tau_2)}}.\end{aligned}$$

If $\tau_2-\tau_1>0$ (that is, $\frac{1+\frac{z\lambda}{z-\delta}\frac{\phi n(\tau_2)^\alpha}{\dot{k}(\tau_2)}}{1-\frac{\lambda}{z-\delta}\phi\alpha n(\tau_2)^{\alpha-1}}<\frac{1}{1-\lambda}$ or $\frac{1}{z-\delta}\phi\alpha n\left(\tau_2\right)^{\alpha-1}+(1-\lambda)\frac{z}{z-\delta}\frac{\phi n(\tau_2)^\alpha}{\dot{k}(\tau_2)}<1$), there are three stages, and if $\tau_2-\tau_1<0$, the second stage does not exist and there is a jump at τ_1 for the $n\left(t\right)$ process. We focus on parameter values such that $\tau_2-\tau_1>0$ in SP.

ii) We show the unique, two-phase intervention to implement SP. As long as the consumption growth $\frac{\dot{n}(t)}{n(t)}$ and $\tau_2-\tau_1$ are implemented, the whole process $\{n\left(t\right)\}$ is implemented, since the lifetime budget constraint of households and $\frac{\dot{n}(t)}{n(t)}$ together determine the initial consumption $n\left(0\right)$ while the lifetime budget constraint (given by the budget constraints plus the transversality conditions) is the same for CE and SP. First, consider stage 3. In CE with subsidy, the $n\left(t\right)$ process is given by $n^{-\sigma}-mz=0$ and $\frac{\dot{m}(t)}{m(t)}=-((z-\delta)\left(1+\mathcal{S}\left(t\right)\right)-\rho)$, which imply $\left.\frac{\dot{n}(t)}{n(t)}\right|_{\text{tax}}=-\frac{1}{\sigma}\frac{\dot{m}(t)}{m(t)}=\frac{(z-\delta)(1+\mathcal{S}(t))-\rho}{\sigma}$. In SP, the $n\left(t\right)$ process is given by $n^{-\sigma}\left(1+\left(\mu\phi\alpha n^{\alpha-1}\right)/\left(n^{-\sigma}\right)\right)-mz=0$ and $\frac{\dot{m}(t)}{m(t)}=-(z-\delta-\rho)$, which imply $\left.\frac{\dot{n}(t)}{n(t)}\right|_{\text{SP}}=-\frac{1}{\sigma}\frac{\dot{m}(t)}{m(t)}+\frac{1}{\sigma}\frac{d\log\left(1+\mu\phi\alpha n^{\alpha-1}/n^{-\sigma}\right)}{dt}$. Note that $n\left(\tau_2\right)=J\left(0\right)$ in both cases and we only need to implement $\frac{\dot{n}(t)}{n(t)}$. It follows $\mathcal{S}\left(t\right)=\frac{1}{z-\delta}\frac{d\log\left(1+\mu\phi\alpha n^{\alpha-1}/n^{-\sigma}\right)}{dt}$. Using $\frac{\mu(t)}{m(t)}=\frac{z\lambda}{z-\delta}$ and $n^{-\sigma}=m_{\text{SP}}z\left(1-\frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}\right)$, we obtain $\mathcal{S}\left(t\right)=\frac{1}{z-\delta}\frac{-d\log\left(1-\frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}\right)}{dt}=\frac{1}{z-\delta}\frac{\frac{\lambda}{z-\delta}\phi\alpha(\alpha-1)n^{\alpha-2}}{1-\frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}}\frac{dn}{dt}$.

Second, consider stage 2. For SP, $\tau_2-\tau_1=\frac{-\log(1-\lambda)}{z-\delta-\rho}+\left(\frac{1}{z-\delta-\rho}\log\frac{n(\tau_1)^{-\sigma}/z}{n(\tau_2)^{-\sigma}/z+\phi\alpha n(\tau_2)^{\alpha-1}\mu(\tau_2)/z+\frac{\phi n(\tau_2)^\alpha}{\dot{k}(\tau_2)}\mu(\tau_2)}\right)$. For CE with subsidy, $n^{-\sigma}-mz\left(1-\lambda\right)=0$ at τ_1, $n^{-\sigma}-mz=0$ at τ_2, and $\frac{\dot{m}(t)}{m(t)}=-((z-\delta)\left(1+\mathcal{S}\left(t\right)\right)-\rho)$ at $t\in[\tau_1,\tau_2]$, which yield $\tau_2-\tau_1=\frac{-\log(1-\lambda)}{(z-\delta)(1+\mathcal{S}(t))-\rho}$. Hence, $\mathcal{S}\left(t\right)=\frac{z-\delta-\rho}{z-\delta}\frac{-\log\vartheta}{\log\vartheta-\log(1-\lambda)}$, where $\vartheta=\frac{n(\tau_1)^{-\sigma}/z}{n(\tau_2)^{-\sigma}/z+\phi\alpha n(\tau_2)^{\alpha-1}\mu(\tau_2)/z+\frac{\phi n(\tau_2)^\alpha}{\dot{k}(\tau_2)}\mu(\tau_2)}$. We can simplify $\vartheta=\frac{n(\tau_1)^{-\sigma}/z}{m^+(\tau_2)+\frac{\phi n(\tau_2)^\alpha}{\dot{k}(\tau_2)}\mu(\tau_2)}=\frac{1-\frac{\lambda}{z-\delta}\phi\alpha n^{\text{SP}}(\tau_2)^{\alpha-1}}{1+\frac{\phi n^{\text{SP}}(\tau_2)^\alpha}{\dot{k}^{\text{SP}}(\tau_2)}\frac{z\lambda}{z-\delta}}$ by using the results of $\frac{\mu(t)}{m(t)}=\frac{z\lambda}{z-\delta}$ and $n^{-\sigma}=m_{\text{SP}}z\left(1-\frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}\right)$ at $t=\tau_2^+$.

Proof of Corollary 2: Consider a policy that subsidizes (taxes) the cost of capital for consumption-goods producers. Specifically, within the consumption goods sector, the subsidy rate on the capital cost of firm j is $s\left(j,t\right)$; that is, firm j only needs to pay the cost $\left(1-s\left(j,t\right)\right)\varphi\left(j|J\left(t\right)\right)$ when it incurs the unit cost $\varphi\left(j|J\left(t\right)\right)$. We assume that, with policy, hierarchical demand for consumption varieties still follows in equilibrium, with the consumption frontier being $n\left(t\right)$. Under this policy, the decentralized household's budget constraint becomes $\dot{k}\left(t\right)=\left(z-\delta\right)k\left(t\right)-\underbrace{\int_0^{n(t)}p\left(j,t\right)c\left(j,t\right)dj}_{=E(t)}-zT\left(t\right)$,

that is, $\dot{k}(t) = (z-\delta)k(t) - \underbrace{R(t)\int_0^{n(t)}(1-s(j,t))\varphi(j|J(t))\,dj}_{=E(t)} - zT(t)$ or

$$\dot{k}(t) = (z-\delta)k(t) - z\underbrace{\left[\overbrace{\int_0^{n(t)}\varphi(j|J(t))\,dj}^{=\mathbf{e}(n(t);J(t))} - \overbrace{\int_0^{n(t)}s(j,t)\varphi(j|J(t))\,dj}^{=\Psi(n(t))}\right]}_{=E(t)} - zT(t)$$

where $T(t) = \Psi(n(t))$ is the lump-sum tax on consumption goods producers.

We rewrite the current value Hamiltonian for CE with subsidy:

$$\mathcal{H}^{CE} = \frac{1}{1-\sigma}n^{1-\sigma} + m\left[(z-\delta)k(t) - z\left(\mathbf{e}(n(t);J(t)) - \Psi(n(t))\right) - zT(t)\right].$$

So the first-order condition with respect to n implies $n^{-\sigma} - mz\left(\mathbf{e}_n(n(t);J(t)) - \Psi'(n(t))\right) = 0$. The SP FOC is $n^{-\sigma} - m_{\text{SP}}z(\mathbf{e}_n(n;J) - \Upsilon(n,J)) = 0$, where $\Upsilon(n,J) \equiv \frac{\mu\mathcal{J}_n(n,J)}{m_{\text{SP}}z}$ is the externality wedge. To implement SP, the two first-order conditions should coincide. That is, $\Psi'(n(t)) = \Upsilon(n(t),J(t))$. Note that $\Psi'(n(t)) = s(n(t),t)\varphi(n(t)|J(t))$. Hence, $s(n(t),t)\varphi(n(t)|J(t)) = \Upsilon(n(t),J(t))$; that is, the instrument must reduce the unit cost of the frontier good n by exactly $\Upsilon(n(t),J(t))$.

i) For stage 3, $\Upsilon(n,J) = \frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}$. Hence, the unit cost for good n should be deducted by $\frac{\lambda}{z-\delta}\phi\alpha n^{\alpha-1}$. To implement $n^{\text{SP}}(t)$ for $t > \tau_2^{\text{SP}}$, subsidize the representative firm in industries $j \in \left[j^{\text{tax}},\, j^{\text{tax}}+\varpi\right]$ with capital $\frac{\lambda}{z-\delta}\phi\alpha j^{\alpha-1}$ when they incur the unit capital cost $\varphi(j|J(t))$. And $n^{\text{SP}}(t)$ needs to belong to $\left[j^{\text{tax}},\, j^{\text{tax}}+\varpi\right]$.

ii) For stage 1, $\frac{\dot{n}}{n}$ is the same for CE and SP. We impose a constant negative subsidy (tax) on the capital cost of consumption goods producers. This policy does not change the growth rate of n in stage 1, but it, together with the policy in stage 3, implement $\tau_2 - \tau_1$. Concretely, in SP, $\frac{m(\tau_1)}{m^-(\tau_2)} = \frac{m(\tau_1)}{m^+(\tau_2)\left[1+\frac{z\lambda}{z-\delta}\frac{\phi n(\tau_2)^\alpha}{(z-\delta)k(\tau_2)-zn(\tau_2)+z\lambda J(\tau_2)}\right]} = \frac{J_0^{-\sigma}/(z(1-\lambda))}{m^+(\tau_2)\left[1+\frac{z\lambda}{z-\delta}\frac{\phi n(\tau_2)^\alpha}{(z-\delta)k(\tau_2)-zn(\tau_2)+z\lambda J(\tau_2)}\right]}$ by using (A.9) and $\frac{\mu(t)}{m(t)} = \frac{z\lambda}{z-\delta}$. This determines $\tau_2-\tau_1$ for SP. The same $\frac{m(\tau_1)}{m^-(\tau_2)}$ must be achieved in CE with subsidy. Since $m(\tau_2)$ is implemented in stage 3, the intervention in stage 1 needs to implement $m(\tau_1) = \frac{J_0^{-\sigma}/(z(1-\lambda))}{1+\frac{z\lambda}{z-\delta}\frac{\phi n(\tau_2)^\alpha}{(z-\delta)k(\tau_2)-zn(\tau_2)+z\lambda J(\tau_2)}}$. The first-order condition in stage 1 with subsidy for CE is $n^{-\sigma} - mz\left[(1-\lambda) - \Psi'(n)\right] = 0$, which implies $m(\tau_1) = \frac{J_0^{-\sigma}/(z(1-\lambda))}{1-\frac{\Psi'(n)}{1-\lambda}}$. Hence, implementing $\frac{J_0^{-\sigma}/(z(1-\lambda))}{1-\frac{\Psi'(n(\tau_1))}{1-\lambda}} = \frac{J_0^{-\sigma}/(z(1-\lambda))}{1+\frac{z\lambda}{z-\delta}\frac{\phi n(\tau_2)^\alpha}{(z-\delta)k(\tau_2)-zn(\tau_2)+z\lambda J(\tau_2)}}$ requires $\Psi'(n) = -\left(\frac{\phi J_0^\alpha}{k^{\text{SP}}(\tau_2)(z-\delta)-z(1-\lambda)n^{\text{SP}}(\tau_2)}\right)\left(\frac{z\lambda}{z-\delta}\right)(1-\lambda)$. This implies that, for $t < \tau_1^{\text{SP}}$, the capital cost of consumption goods producers $j \in \left[j^{\text{tax}},\, j^{\text{tax}}+\varpi\right]$

needs to be increased by $\left(\frac{\phi J_0^{\alpha}}{k^{\mathrm{SP}}(\tau_2)(z-\delta)-z(1-\lambda)n^{\mathrm{SP}}(\tau_2)}\right)\left(\frac{z\lambda}{z-\delta}\right)(1-\lambda)$ when they incur the capital cost $1-\lambda$. And $n^{\mathrm{SP}}(t)$ needs to belong to $\left[j^{\mathrm{tax}},\, j^{\mathrm{tax}}+\varpi\right]$. The switching time from the negative to the positive subsidy can be any $t^{*\mathrm{tax}} \in \left[\tau_1^{\mathrm{SP}}, \tau_2^{\mathrm{SP}}\right]$.

Proof of Proposition 3: We compare the competitive equilibrium with the planner's solution. For convenience, we denote $J(t)$ by J_t. First consider the model without spillover. We can find a t^* large enough and a Δ small enough such that $J_{t^*}^{\mathrm{SP}} - J_{t^*}^{\mathrm{CE}} > \Delta$. At t^*, we introduce the spillover specification as follows: Letting $\bar{J} = J_{t^*}^{\mathrm{CE}} + \Delta$, for $t > t^*$,

$$z(t) = \begin{cases} z\psi(J(t)) & \text{if } J_{t^*} < \bar{J} \\ z(1+\epsilon)\psi(J(t)) & \text{if } J_{t^*} \geq \bar{J} \end{cases}, \tag{A.10}$$

where ϵ is a small positive number, and $J_{t^*}^{\mathrm{CE}}$ is the J_t level at $t = t^*$ for the competitive equilibrium without spillover—we can write this $J_{t^*}^{\mathrm{CE}}$ as $J_{t^*}^{\mathrm{CE}}(\epsilon = 0)$.

With the spillover specification given above, we now look at the competitive equilibrium first. We show that under the sufficient condition that ϵ is enough, even after agents look forward and take account of the possible regime change given in (A.10), the revised policy for the competitive equilibrium results in J_{t^*} still being lower than $\bar{J} = J_{t^*}^{\mathrm{CE}} + \Delta$, so a jump in $z(t)$ at $t = t^*$ will not happen. Concretely, given the specification of $z(t)$ in (A.10), agents will re-optimize starting from their initial period. The result is that ϵ may introduce a deviation in their individual policy function along the transition path from their original policy (which corresponds to $\epsilon = 0$). By the continuity of policies, for any given small enough ϵ, there exists $\xi < \infty$ such that the wedge (gap) in the equilibrium J_t process resulted from the deviation in the policy is bounded by $\xi\epsilon$, that is, $\mid J_t^{\mathrm{CE}}(\epsilon) - J_t^{\mathrm{CE}}(\epsilon = 0) \mid < \xi\epsilon$, for $\forall t \leq t^*$. Then, the competitive equilibrium J_t will not exceed the threshold at t^* if ϵ is small enough, that is, $J_{t^*}^{\mathrm{CE}}(\epsilon) - J_{t^*}^{\mathrm{CE}}(\epsilon = 0) \leq \mid J_t^{\mathrm{CE}}(\epsilon) - J_t^{\mathrm{CE}}(\epsilon = 0) \mid < \xi\epsilon < \Delta$. Therefore, the competitive equilibrium does not pass the threshold. For the social planner's equilibrium, it passes the threshold even if $\epsilon = 0$. Therefore, for $t > t^*$, $z^{\mathrm{SP}}(t) = z(1+\epsilon)\psi(J(t))$ and $z^{\mathrm{CE}}(t) = z\psi(J(t))$.

B Additional Results

Results in Section 5.2: First, we derive $c(j,t)$. As in Bertola, Foellmi and Zweimuller (2014, Chapter 12), if $\theta + \gamma > 0$, there exists a cutoff variety $n(t)$. Once we know $n(t)$, we can find $c(j,t)$ by equalling the marginal utility-price ratio across $j \in (0, n(t)]$. We have $c(j,t) = s\left(1 - \frac{\varphi(j|J(t))j^{\gamma+\theta}}{\varphi(n(t)|J(t))n(t)^{\gamma+\theta}}\right)$ for $j \in (0, n(t)]$.

Next, we derive the cost function $\mathbf{e}(n(t); J(t)) = \int_0^{n(t)} c(j,t) \cdot \hat{\varphi}(j|J(t))\, dj$. Substituting $c(j,t)$

yields

$$\mathbf{e}(n;J)/s=\int_0^n\hat{\varphi}(j|J)\left(1-\frac{\hat{\varphi}(j|J)}{\hat{\varphi}(n|J)}\frac{j^{\gamma}}{n^{\gamma}}\right)dj=\left(\int_0^{\tilde{n}}\tilde{\varphi}\left(\tilde{j}\right)d\tilde{j}-\tilde{\varphi}(\tilde{n})^{-1}\tilde{n}^{-\gamma}\int_0^{\tilde{n}}\tilde{\varphi}\left(\tilde{j}\right)^2\tilde{j}^{\gamma}d\tilde{j}\right)J^{\theta+1},$$

where $\tilde{n}=n/J$, $\tilde{j}=j/J$ and $\tilde{\varphi}\left(\tilde{j}\right)\equiv\tilde{j}^{\theta}\left[1-\lambda\left(1+\kappa\varepsilon\tilde{j}^{-\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{1}{\varepsilon-1}}\right]$. Plugging the functional forms of $\tilde{\varphi}\left(\tilde{j}\right)$ and defining a new variable $u\equiv\tilde{j}^{-\frac{\varepsilon-1}{\varepsilon}}$, the above integration can be written as

$$\begin{aligned}\mathbf{e}(n;J)/s&=\frac{\varepsilon}{1-\varepsilon}J^{\theta+1}\left\{\begin{array}{l}\int_0^{\tilde{n}^{\frac{1-\varepsilon}{\varepsilon}}}u^{\frac{\varepsilon(\theta+1)}{1-\varepsilon}-1}\left[1-\lambda\left(1+\kappa\varepsilon u\right)^{\frac{1}{\varepsilon-1}}\right]du\\-\tilde{\varphi}(\tilde{n})^{-1}\tilde{n}^{-\gamma}\int_0^{\tilde{n}^{\frac{1-\varepsilon}{\varepsilon}}}u^{\frac{\varepsilon(2\theta+\gamma+1)}{1-\varepsilon}-1}\left[1-\lambda\left(1+\kappa\varepsilon u\right)^{\frac{1}{\varepsilon-1}}\right]du\end{array}\right\}\\&=\left[\frac{\tilde{n}^{1+\theta}}{1+\theta}h_1\left(\tilde{n};\frac{\varepsilon(\theta+1)}{1-\varepsilon}-1\right)-\tilde{\varphi}(\tilde{n})^{-1}\tilde{n}^{-\gamma}\frac{\tilde{n}^{2\theta+\gamma+1}}{2\theta+\gamma+1}h_2\left(\tilde{n};\frac{\varepsilon(2\theta+\gamma+1)}{1-\varepsilon}-1\right)\right]J^{\theta+1}\\&=\left[h_1\left(\frac{n}{J};\frac{\varepsilon(\theta+1)}{1-\varepsilon}-1\right)-\frac{1+\theta}{(1+\theta)+(\gamma+\theta)}h_2\left(\frac{n}{J};\frac{\varepsilon(2\theta+\gamma+1)}{1-\varepsilon}-1\right)\right]\frac{n^{1+\theta}}{1+\theta}.\end{aligned}$$

where

$$\begin{aligned}h_1(x;\nu)&=1-\lambda\mathrm{H}\left(\frac{1}{1-\varepsilon},\nu+1;\nu+2;-\kappa\varepsilon x^{\frac{1-\varepsilon}{\varepsilon}}\right)\\h_2(x;\nu)&=\left[\begin{array}{c}1-2\lambda\mathrm{H}\left(\frac{1}{1-\varepsilon},\nu+1;\nu+2;-\kappa\varepsilon x^{\frac{1-\varepsilon}{\varepsilon}}\right)\\+\lambda^2\mathrm{H}\left(\frac{2}{1-\varepsilon},\nu+1;\nu+2;-\kappa\varepsilon x^{\frac{1-\varepsilon}{\varepsilon}}\right)\end{array}\right]\left[1-\lambda\left(1+\kappa\varepsilon\left(\frac{n}{J}\right)^{-\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{1}{\varepsilon-1}}\right]^{-1}\end{aligned}$$

with H being the hypergeometric function, defined as:

$$\mathrm{H}(a,b,c;x)=\frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{n=0}^{\infty}\frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)}\frac{x^n}{n!},$$

where $\Gamma(x)$ is the Gamma function. In the above derivation, we use the following two equalities:

$$\begin{aligned}\int_0^{x^{\frac{1-\varepsilon}{\varepsilon}}}u^{\nu}\left[1-\lambda(1+\kappa\varepsilon u)^{\frac{1}{1-\varepsilon}}\right]du&=\frac{x^{(1+\nu)\frac{1-\varepsilon}{\varepsilon}}}{1+\nu}1-\lambda\mathrm{H}\left(\frac{1}{1-\varepsilon},\nu+1;\nu+2;-\kappa\varepsilon x^{\frac{1-\varepsilon}{\varepsilon}}\right)\\\int_0^{x^{\frac{1-\varepsilon}{\varepsilon}}}u^{\nu}\left[1-\lambda\left(1+\kappa\varepsilon u\right)^{\frac{1}{\varepsilon-1}}\right]^2du&=\frac{x^{(1+\nu)\frac{1-\varepsilon}{\varepsilon}}}{1+\nu}\left[\begin{array}{c}1-2\lambda\mathrm{H}\left(\frac{1}{1-\varepsilon},\nu+1;\nu+2;-\kappa\varepsilon x^{\frac{1-\varepsilon}{\varepsilon}}\right)\\+\lambda^2\mathrm{H}\left(\frac{2}{1-\varepsilon},\nu+1;\nu+2;-\kappa\varepsilon x^{\frac{1-\varepsilon}{\varepsilon}}\right)\end{array}\right].\end{aligned}$$

Note in the special case of $s=1$ and $\theta=0$, the term $\frac{n^{1+\theta}}{1+\theta}h_1\left(\frac{n}{J};\frac{\varepsilon(\theta+1)}{1-\varepsilon}-1\right)$ collapses into the CES form:

$$\begin{aligned}&h_1\left(\frac{n}{J};\frac{\varepsilon}{1-\varepsilon}-1\right)n=n\left[1-\lambda\mathrm{H}\left(\frac{1}{1-\varepsilon},\frac{\varepsilon}{1-\varepsilon},\frac{1}{1-\varepsilon},-\kappa\varepsilon\left(\frac{n}{J}\right)^{\frac{1-\varepsilon}{\varepsilon}}\right)\right]\\&=n-\lambda\left(n^{\frac{\varepsilon-1}{\varepsilon}}+(\kappa\varepsilon)J^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}},\end{aligned}$$

where we note $\mathrm{H}\left(\frac{1}{1-\varepsilon},\frac{\varepsilon}{1-\varepsilon},\frac{1}{1-\varepsilon},-\kappa\varepsilon\left(\frac{n}{J}\right)^{\frac{1-\varepsilon}{\varepsilon}}\right)=\left(1+\kappa\varepsilon\left(\frac{n}{J}\right)^{\frac{1-\varepsilon}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$.

Finally, we derive the formula for $c(t)\equiv U(n(t),J(t))$. Plugging $c(j,t)$ into $c(t)$ yields

$$\begin{aligned}U(n,J)&=\frac{s^2}{2}\int_{\Delta}^{n}j^{-\gamma}dj-\frac{s^2}{2}\int_{0}^{n}j^{-\gamma}\left(\frac{\hat{\varphi}(j|J)}{\hat{\varphi}(n|J)}\frac{j^{\gamma}}{n^{\gamma}}\right)^{2}dj=\frac{s^2}{2}\left\{\frac{n^{1-\gamma}-\Delta^{1-\gamma}}{1-\gamma}-J^{1-\gamma}\int_{0}^{\tilde{n}}\tilde{j}^{-\gamma}\left(\frac{\tilde{\varphi}\left(\tilde{j}\right)}{\tilde{\varphi}\left(\tilde{n}\right)}\frac{\tilde{j}^{\gamma}}{\tilde{n}^{\gamma}}\right)^{2}d\tilde{j}\right\}\\&=\frac{s^2}{2}\left\{\frac{n^{1-\gamma}-\Delta^{1-\gamma}}{1-\gamma}-J^{1-\gamma}\tilde{\varphi}\left(\tilde{n}\right)^{-2}\tilde{n}^{-2\gamma}\left[\int_{0}^{\tilde{n}}\tilde{j}^{2\theta+\gamma}\left[1-\lambda\left(1+\kappa\varepsilon\tilde{j}^{-\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{1}{\varepsilon-1}}\right]^{2}d\tilde{j}\right]\right\}\\&=\frac{1}{2}s^2\left[1-\frac{\gamma-1}{(\gamma+1)+(\gamma+\theta)}\frac{1}{(n/\Delta)^{\gamma-1}-1}h_3\left(\frac{n}{J};\frac{\varepsilon(2\theta+\gamma+1)}{1-\varepsilon}-1\right)\right]\frac{\Delta^{1-\gamma}-n^{1-\gamma}}{\gamma-1}\end{aligned}$$

where $h_3(x;\nu)$ is defined as

$$h_3(x;\nu)=\left[\begin{array}{l}1-2\lambda\mathrm{H}\left(\frac{1}{1-\varepsilon},\nu+1,\nu+2;-\kappa\varepsilon x^{\frac{1-\varepsilon}{\varepsilon}}\right)\\+\lambda^2\mathrm{H}\left(\frac{2}{1-\varepsilon},\nu+1,\nu+2;-\kappa\varepsilon x^{\frac{1-\varepsilon}{\varepsilon}}\right)\end{array}\right]\left[1-\lambda\left(1+\kappa\varepsilon x^{-\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{1}{\varepsilon-1}}\right]^{-2}.$$

Calibration in Section 5.3: For the externally assigned parameters, we use exactly the same ones as in Table 4, except for two new ones: L and $\hat{\alpha}$. These two parameters, along with the internally calibrated ones, are listed in the following table.

Table B-1: Model parameterization

Parameter	Symbol	Value	Source/Notes
externally assigned parameters			
labor endowment	L	1.000	normalized
capital share	$\hat{\alpha}$	0.333	Acemoglu (2009)
internally calibrated parameters			
experience accumulation (later)	ϕ	0.0433	–
capital accumulation	z	0.2043	–
initial capital	K_0	0.1072	–
experience accumulation (curvature)	ζ	0.0722	–

Table B-2 reports the moment matching results.

Table B-2: Moments

Moment	Model	Data	Data Source
years to rise of consumption	29.300	25.000	Matsuyama (2002)
market size impact on productivity	0.509	0.470	our regression result
growth of K (max)	0.134	0.070	Barro (2003)
growth of K (average)	0.105	0.057	Barro (2003)
growth of N (max)	0.028	0.057	Greenwood and Uysal (2005)
growth of N (average)	0.026	0.036	Greenwood and Uysal (2005)