Online Appendix
The Stable Transformation Path

Francisco J. Buera* Joseph Kaboski† Martí Mestieri‡
Daniel G. O'Connor§

August 17, 2020

There are four sections in this appendix. Appendix A contains the proof of Theorem 1 in the main text. Appendix B shows how the models of Section 2.1 and 4.4 map as special cases of the more general model of Section 3. We also show how the Acemoglu and Guerrieri (2008) model is nested in the general model. Appendix C providea details of data construction for the empirical aspects of the paper in Section 4. Appendix D presents detailed derivations for the nonhomothetic CES variant of the model presented in Section 4.4.

A Proof of Existence and Uniqueness of the STraP

This section provides a proof of Theorem 1 in the main text. The proof consists of two steps. The first step is standard, and we provide a succinct description (referring the reader to Acemoglu (2009) for details). The second step is novel, and we devote the majority of the proof to provide a detailed derivation.

The first step in proving our result is to characterize the solution of the Planner's problem going forward given capital k at time t. We begin noting that the assumptions on f and u guarantee that the planner's solution is interior. Moreover, Theorem 7.9 in Acemoglu (2009) holds, so any solution must satisfy the first order conditions of the Hamiltonian

$$H(t,k,c,\lambda) = e^{-\rho t}\mathcal{A}_u(t)u(c(t),t) + \lambda(t)\left[f(k(t),t) - (\delta+\gamma_k(t))\,k(t) - c(t)\right],$$

which yield the the Euler Equation

$$\frac{\dot{c}(t)}{c(t)} = -\frac{\frac{\partial u(c(t),t)}{\partial c}}{\frac{\partial^2 u(c(t),t)}{\partial c^2}c(t)}\left[\frac{\partial f(k(t),t)}{\partial k} - \delta - \gamma_k(t) - \rho + \gamma_u(t)\right]. \tag{A.1}$$

*Washington University in St. Louis. Email: fjbuera@wstlu.edu
†University of Notre Dame. Email: jkaboski@nd.edu
‡Federal Reserve Bank of Chicago. Email: marti.mestieri@gmail.com
§Massachusetts Institute of Technology. Email: doconn@mit.edu

The system is further characterized by the law of motion

$$\dot{k}(t) = f(k(t), t) - (\delta + \gamma_k(t))\, k(t) - c(t). \tag{A.2}$$

Theorem 7.12 in Acemoglu (2009) holds (since the value function is differentiable and its derivative converges to 0 along any feasible path), implying that any interior solution path satisfies the transversality condition

$$\lim_{t\to\infty} H(t, k(t), c(t), \lambda(t)) = 0. \tag{A.3}$$

Finally, the maximized Hamiltonian is strictly concave in the state variable so the continuous path defined by the law of motion (equation A.2), the Euler equation (equation A.1), and the transversality condition (equation A.3) is the unique solution to the planner's problem.

Now we proceed to the second step of the proof. In this part, we need to ensure that the system is Lipschitz continuous to invoke some of the results from the theory of ordinary differential equations. This requires that f itself is Lipschitz continuous. We therefore leverage on the Inada condition assumption for f, and define an $\underline{\varepsilon} > 0$ that bounds capital away from 0 with the property that if $k(\tau) \in [\underline{\varepsilon}, \bar{k}]$, then $k(t) \in [\underline{\varepsilon}, \bar{k}]$ for all $t > \tau$. Given our assumption that $f \in C^2$ for all $(k, t) \in (0, \infty) \times \mathbb{R}$, f is Lipschitz continuous in $[\underline{\varepsilon}, \bar{k}]$ and we focus on that interval. We also define a $\bar{C}$ so that for all $t \in \mathbb{R}$ and $k \in [\underline{\varepsilon}, \bar{k}]$, the optimal consumption $c(t) \in (0, \bar{C})$. Then, we can restrict the domain of consumption to the compact interval $[0, \bar{C}]$ without loss.

The Euler equation and the law of motion for capital determine a two dimensional non-autonomous system. In the first step of our proof, we have shown that there is a unique consumption level for a given starting capital k and time t consistent with optimization. This function, which we label $c(k, t)$, is the unique consumption level that shoots to the asymptotic balanced growth path with capital level k_∞. Using this observation, we can write our system as a one dimensional non-autonomous differential equation

$$\dot{k}(t) = f(k(t), t) - (\delta + \gamma_k(t))\, k(t) - c(k(t), t). \tag{A.4}$$

By construction, given initial conditions $(k_0, \tau) \in [\underline{\varepsilon}, \bar{k}] \times \mathbb{R}$, $k(t) \to k_\infty$ as $t \to \infty$. Proving existence and uniqueness of the STraP comes down to proving that there exists one unique path of this system that has $k(t) \to k_{-\infty}$ as $t \to -\infty$. To do that, we will use the anti-funnel existence and uniqueness Theorem 4.7.5 from Hubbard and West (1991).

Let us outline how the rest of the proof proceeds. At a broad level, the main complication in this step is characterizing $c(k, t)$. Define $c_-(k)$ as the optimal choice of consumption given a starting capital amount in the negative asymptotic optimal growth problem. Define $c_+(k)$ similarly for the positive asymptotic optimal growth problem. The usual growth phase plane analysis gives us information about $c_-(k)$ and $c_+(k)$. For example, linearizing $c_-(k)$ around $k_{-\infty}$ tells us that c is continuously differentiable and increasing in k around the steady state. In Lemma 1, we prove that under the conditions stated in Assumption 1, these properties extend to $c(k,t)$ for t close enough to $-\infty$. We use this result to check some of conditions needed for the anti-funnel theorem in the ensuing lemmas. In Lemma 3, we show that equation (A.4) gets close to the law of motion for the negative asymptotic optimal growth problem for $t \to -\infty$. Then using these results, we construct an anti-funnel and apply the anti-funnel theorem in Proposition 1.

We begin using Assumption 1 to characterize the asymptotic behavior of $c(k,t)$. Assumption 1 allows us to reparametrize time to lie in the compact interval $[0,1]$ where 0 corresponds to $t=-\infty$ and 1 corresponds to $t=\infty$. Then we can use continuity to extend results about functions at $t=\pm\infty$ to functions where $t \in \mathbb{R}$. The assumptions on h, h', and h'' guarantee that the system of differential equations remain well behaved as $t \to \pm\infty$. In particular, we need that $\dot{\gamma}_k(t)$ and $\dot{\gamma}_u(t)$ approach 0 fast enough so that the differential system at sufficiently large and small t looks like the asymptotic differential systems. This allows us to linearize around the asymptotic steady states and gain information about the behavior near them.

Lemma 1. *Suppose that Assumption 1 holds. Then, there exists a function* $\check{c} : [\underline{\varepsilon}, \bar{k}] \times [0,1] \to [0, \bar{C}]$ *with the following properties:*

- $\check{c}(k,0) = c_-(k)$ *for all* $k \in [\underline{\varepsilon}, \bar{k}]$,
- $\check{c}(k,1) = c_+(k)$ *for all* $k \in [\underline{\varepsilon}, \bar{k}]$,
- $\check{c}(k,z) = c(k,h(z))$ *for all* $k \in [\underline{\varepsilon}, \bar{k}]$ *and* $z \in (0,1)$, *and*
- $\check{c}(k,z)$ *is continuously differentiable.*

Proof. We can rewrite our non-autonomous 2 dimensional differential system as a 3 dimensional autonomous system by including time as a variable which we will write as $s(t)$. Then the system is

$$\dot{k}(t) = f(k(t),s(t)) - (\delta + \gamma_k(s(t)))\,k(t) - c(t), \tag{A.5}$$

$$\dot{c}(t) = \frac{c(t)}{\theta(c(t),s(t))}\left[\frac{\partial f(k(t),s(t))}{\partial k} - \delta - \gamma_k(s(t)) - \rho + \gamma_u(s(t))\right], \tag{A.6}$$

$$\dot{s}(t) = 1. \tag{A.7}$$

The function $s(t)$ is unbounded in this system, which makes it cumbersome to deal with. Therefore, we reparametrize time with $z = h(s) \in (0,1)$. This autonomous system, dropping the explicit time notation, can be written as

$$\dot{k} = f(k,h^{-1}(z)) - \left(\delta + \gamma_k(h^{-1}(z))\right)k - c, \tag{A.8}$$

$$\dot{c} = \frac{c}{\theta(c,h^{-1}(z))}\left[\frac{\partial f(k,h^{-1}(z))}{\partial k} - \delta - \gamma_k(h^{-1}(z)) - \rho + \gamma_u(h^{-1}(z))\right], \tag{A.9}$$

$$\dot{z} = h'(h^{-1}(z)), \tag{A.10}$$

and it is defined on the box $[\varepsilon, \bar{k}] \times [0, \bar{C}] \times (0, 1)$. We denote the right hand side of equations (A.8), (A.9) and (A.10) by $F_k(k, c, z)$, $F_c(k, c, z)$, and $F_z(k, c, z)$, respectively. The system can easily be extended to the closed box $[\varepsilon, \bar{k}] \times [0, \bar{C}] \times [0, 1]$ by replacing $\gamma_k(h^{-1}(z))$, $\gamma_u(h^{-1}(z))$, $f(k, h^{-1}(z))$, $\frac{\partial f(k,h^{-1}(z))}{\partial k}$, $\theta(c, h^{-1}(z))$ and $h'(h^{-1}(z))$ with their limits as $z \to 0, 1$. The limits of the γ terms, f, and θ follow immediately from our assumptions. Since $h(t)$ converges to a real number and h' converges, it follows that h' must converge to 0 as $t \to \pm\infty$.

While the system is defined, it might not be well behaved as z approaches 0 or 1. In particular, the derivative of the functions F_k, F_c, and F_z at those points might not exist and so the system might not be Lipschitz continuous. The total differentials of these functions in $z \in (0, 1)$ are

$$\begin{aligned} dF_k(k, c, z) = &\left[\frac{\partial f(k, h^{-1}(z))}{\partial k} - \delta - \gamma_k(h^{-1}(z))\right] dk - dc \\ &+ \left[\frac{\partial f(k, h^{-1}(z))}{\partial t} - \dot{\gamma}_k(h^{-1}(z))k\right] \left(h^{-1}\right)'(z)dz, \end{aligned} \tag{A.11}$$

$$\begin{aligned} dF_c(k, c, z) = &\frac{c}{\theta(c, h^{-1}(z))} \frac{\partial^2 f(k, h^{-1}(z))}{\partial k^2} dk + F_c(k, c, z) \frac{\theta(c, h^{-1}(z)) - c\frac{\partial \theta(c,h^{-1}(z))}{\partial c}}{c\theta(c, h^{-1}(z))} dc \\ &+ \left\{ \frac{c}{\theta(c, h^{-1}(z))} \left[\frac{\partial^2 f(k, h^{-1}(z))}{\partial k \partial t} + \dot{\gamma}_u(h^{-1}(z)) - \dot{\gamma}_k(h^{-1}(z))\right] (h^{-1})'(z) \right. \\ &\left. - \frac{\frac{\partial \theta(c,h^{-1}(z))}{\partial t}(h^{-1})(z)}{\theta(c, h^{-1}(z))} F_c(k, c, z) \right\} dz, \end{aligned} \tag{A.12}$$

$$dF_z(k, c, z) = h''(h^{-1}(z))(h^{-1})'(z)dz. \tag{A.13}$$

Next, we note that the Inverse Function Theorem together with Assumption 1 guarantee that the derivative converges as $z \to 0, 1$. Therefore, the derivatives at $z = 0, 1$ exist and correspond with the limits. Along with the fact that k is bounded away from 0, we thus have that the system is Lipschitz continuous.

Notice that there are two steady states with positive consumption in the system defined by equations (A.8), (A.9) and (A.10): one at $(k_\infty, c_\infty, 1)$ and another at $(k_{-\infty}, c_{-\infty}, 0)$. Next, we linearize this system of equations around the $(k_\infty, c_\infty, 1)$ steady state. Then a deviation $k_\infty + \hat{k}(t)$, $c_\infty + \hat{c}(t)$, and $1 + \hat{z}(t)$ must locally satisfy

$$\begin{pmatrix} \dot{\hat{k}}(t) \\ \dot{\hat{c}}(t) \\ \dot{\hat{z}}(t) \end{pmatrix} = \begin{pmatrix} \rho - \gamma_{+u} & -1 & 0 \\ \frac{c_\infty}{\theta_+(c_\infty)} f''_+ k_\infty) & 0 & 0 \\ 0 & 0 & a_+ \end{pmatrix} \begin{pmatrix} \hat{k}(t) \\ \hat{c}(t) \\ \hat{z}(t) \end{pmatrix}. \tag{A.14}$$

The associated characteristic polynomial to (A.14) is

$$(a_+ - \lambda) \left(\lambda^2 + (\gamma_{+u} - \rho)\lambda + \frac{c_\infty}{\theta(c_\infty)} f''_+(k_\infty)\right).$$

The system has one positive and two negative eigenvalues. The two negative ones are a_+ with associated eigenvector $(0,0,1)^T$ and $\frac{1}{2}\left(\rho-\gamma_{+u}-\sqrt{(\gamma_{+u}-\rho)^2-4\frac{c_\infty}{\theta_+(c_\infty)}f''_+(k_\infty)}\right)$ with associated eigenvector $\left(1,\frac{1}{2}\left(\rho-\gamma_{+u}+\sqrt{(\gamma_{+u}-\rho)^2-4\frac{c_\infty}{\theta_+(c_\infty)}f''_+(k_\infty)}\right),0\right)^T$. Thus, locally this stable plane defines consumption as a function of k and z.

We can then extend this to a manifold in the box $[\varepsilon,\bar{k}]\times[0,\bar{C}]\times[0,1]$. This manifold will be our function $\check{c}(k,z)$. Notice that setting $z=1$, the system is exactly the differential equations describing the asymptotic growth problem so the manifold corresponds with the graph of $c_+(k)$ trivially. Next we turn to $z\in(0,1)$. For any t and k we know that there exists $c(k,t)$ that converges to the steady state. Therefore, the point $(k,c(k,h(z)),z)$ is on the manifold. If there were another point on the manifold for a given k and z, with that capital and time, starting with either consumption would satisfy the transversality condition and the differential equations. That contradicts uniqueness of the optimal consumption. Therefore, this manifold defines a function $\check{c}(k,z)$ on $[\varepsilon,\bar{k}]\times(0,1]$. We then turn to extending to $z=0$. We will show that the limit of $\check{c}(k,z)$ as $z\to 0$ is $c_-(k)$. Then we can define $\check{c}(k,0)=c_-(k)$. As F_k, F_c, and F_z are continuously differentiable, it follows that $\check{c}(k,z)$ is continuously differentiable.

We now show that $c(k,z)\to c_-(k)$ as $z\to 0$. Starting at time $-\tau$, the time path of consumption solves the problem

$$\max_{c(t),k(t)}\int_{t=0}^{\infty}e^{-\rho t}\frac{\mathcal{A}_u(t-\tau)}{\mathcal{A}_u(-\tau)}u(c(t),t)dt \tag{A.15}$$

such that

$$\dot{k}(t)=f(k(t),t-\tau)-(\delta+\gamma_k(t-\tau))k(t)-c(t) \tag{A.16}$$

and

$$k(0)=k_0. \tag{A.17}$$

where we have simply divided the original expression by $A_u(-\tau)e^{-\rho\tau}$. Replacing τ with $h^{-1}(z)$ makes it a maximization problem with two parameters, $k_0\in\left[\underline{\varepsilon},\bar{k}\right]$ and $z\in[0,1]$. When $z=0,1$ the problem is replaced by the asymptotic growth problems.

Our next step is to apply Berge's maximum theorem. In order to apply it, we need to ensure that the utility function is continuous and the constraint correspondence is continuous, compact valued, and contains no empty values. Using the discounted norm $||c(\cdot)||\equiv\int_{t=0}^{\infty}e^{-\rho t}c(t)dt$ this is the case for all $z\in(0,1)$. We just need to ensure that the functions are continuous at $z=0,1$. Here we focus on $z=0$ (an analogous argument can be done for $z=1$). We use the Dominated Convergence Theorem to prove this result. We have that $\frac{\mathcal{A}_u(t-\tau)}{\mathcal{A}_u(-\tau)}=e^{\int_0^t\gamma_u(s-\tau)ds}$, where $-\tau=h^{-1}(z)$. Take $\gamma_u>\max\{\gamma_{+u},\gamma_{-u}\}$ but $\gamma_u<\rho$. Then since $\gamma_u(t)$ converges on either side and it is twice-continuously differentiable, there is a B so that $\frac{\mathcal{A}_u(t-\tau)}{\mathcal{A}_u(-\tau)}\leq Be^{\gamma_u t}$ for all τ. Since c is also bounded, the absolute value of the function $e^{-\rho t}\frac{\mathcal{A}_u(t-\tau)}{\mathcal{A}_u(-\tau)}u(c(t),t-\tau)$ is bounded above by $e^{(\gamma_u-\rho)t}B\bar{u}$. Thus, by the DCT, we have that $\lim_{z\to 0}\int_0^\infty e^{-\rho t}\frac{\mathcal{A}_u(t+h^{-1}(z))}{\mathcal{A}_u(h^{-1}(z))}u(c(t),t+h^{-1}(z))dt = \int_0^\infty \lim_{z\to 0} e^{-\rho t}\frac{\mathcal{A}_u(t+h^{-1}(z))}{\mathcal{A}_u(h^{-1}(z))}u(c(t),t+h^{-1}(z))dt$. Note that $\lim_{z\to 0} u(c(t),t+h^{-1}(z)) = u_-(c(t))$ and $\lim_{z\to 0}\frac{\mathcal{A}_u(t+h^{-1}(z))}{\mathcal{A}_u(h^{-1}(z))} = e^{\gamma_{-u}t}$. Thus, the utility function is continuous on all of $[\underline{\varepsilon},\bar{k}]\times[0,\bar{C}]\times[0,1]$. By Berge's maximum theorem, optimal consumption is continuous in z and $c(k,z)$ must approach $c_-(k)$ as $z\to 0$. □

We next present a series of Lemmas that build on Lemma 1 that allow us to show that the conditions for Theorem 4.7.5 of Hubbard and West (1991) hold in this setting.[1] To use the anti-funnel theorem in Hubbard and West, we will look at the backwards differential equation. Define $\tilde{k}(t) = k(-t)$. This follows the differential equation

$$\dot{\tilde{k}}(t) = -f\left(\tilde{k}(t),-t\right) + (\delta+\gamma_k(-t))\,\tilde{k}(t) + c\left(\tilde{k}(t),-t\right). \tag{A.18}$$

We then want to show that there exists a unique time path $\tilde{k}(t)$ with $\tilde{k}(t)\to k_{-\infty}$ as $t\to\infty$. Define $F\left(\tilde{k},t\right) \equiv -f\left(\tilde{k}(t),-t\right) + (\delta+\gamma_k(-t))\,\tilde{k}(t) + c\left(\tilde{k}(t),-t\right)$. We similarly define the function that describes the backward capital motion in the negative asymptotic growth problem: $F_-\left(\tilde{k}\right) \equiv -f\left(\tilde{k}(t)\right) + (\delta+\gamma_{-k})\,\tilde{k}(t) + c_-\left(\tilde{k}(t)\right)$.

Lemma 2. *On some interval $U\subset[0,\bar{k}]$ containing $k_{-\infty}$ in its interior, for all $\varepsilon>0$, there exists a $T>0$ so that if $t>T$,*

$$\left|F\left(\tilde{k},-t\right) - F_-\left(\tilde{k}\right)\right| < \varepsilon, \qquad \forall \tilde{k}\in I.$$

Proof. By Lemma 1, $\check{c}(k,z)$ is continuous on the compact region $[\varepsilon,\bar{k}]\times[0,1]$. Therefore, it is absolutely continuous. Then $\forall\varepsilon>0$, there exists a $\delta>0$ such that $|\check{c}(k_1,z_1)-\check{c}(k_2,z_2)|<\varepsilon$, if $\max\{|k_1-k_2|,|z_1-z_2|\}<\delta$, where we use the sup norm. In particular, if $z_2=0$ and $k_1=k_2\equiv k$, then for $z_1<\delta$, $|\check{c}(k,z_1)-\check{c}(k,0)|<\varepsilon$. Transforming into time, if $t_1<h(\delta)$ then

$$|c(k,t_1)-c_-(k)|<\varepsilon, \forall k\in[\varepsilon,\bar{k}].$$

[1] For completeness, we state Theorem 4.7.5 from Hubbard and West (1991) here.

Theorem 1. *Let $\alpha(t)$ and $\beta(t)$, $\beta(t)\le\alpha(t)$, be two fences defined for $t\in[a,b)$ that bound an antifunnel for the differential equation $x'=f(t,x)$. Let $f(t,x)$ satisfy a Lipschitz condition in the antifunnel. Furthermore, let the antifunnel be narrowing, with*

$$\lim_{t\to b}(\alpha(t)-\beta(t)) = 0.$$

If $(\partial f/\partial x)(t,x)\ge w(t)$ in the antifunnel, where $w(t)$ is a function satisfying

$$\int_a^b w(s)ds > -\infty,$$

then there is a unique solution which stays in the antifunnel.

Similarly, there exists a T_2 so that if $t < T_2$, then

$$|f(k,t) - f_-(k)| < \varepsilon.$$

To finish the proof, we simply note that

$$\begin{aligned}\left|F\left(\tilde{k},-t\right) - F_-\left(\tilde{k}\right)\right| &= \left|-f\left(\tilde{k},-t\right) + f_-\left(\tilde{k}\right) + c\left(\tilde{k},-t\right) - c_-\left(\tilde{k}\right) + \gamma_k(-t)\tilde{k} - \gamma_{-k}\tilde{k}\right| \\ &\leq 2\varepsilon + |\gamma_k(-t) - \gamma_{-k}|\,\tilde{k} \\ &\leq 3\varepsilon\end{aligned}$$

for t sufficiently large that $|\gamma_k(-t) - \gamma_{-k}| < \frac{\varepsilon}{k}$. □

Lemma 3. *$F\left(\tilde{k},t\right)$ is Lipschitz continuous on some interval U containing $k_{-\infty}$ in its interior. That is, there exists a $\mathcal{K} > 0$ such that for all $t \in [0,\infty)$ and $\tilde{k}_1, \tilde{k}_2 \in I$,*

$$\left|F\left(\tilde{k}_1,t\right) - F\left(\tilde{k}_2,t\right)\right| \leq \mathcal{K}\left|\tilde{k}_1 - \tilde{k}_2\right|.$$

Proof. Notice that $\left|\frac{\partial f(k,h^{-1}(z))}{\partial k}\right|$ is continuous on the compact interval $[\varepsilon,\bar{k}]\times[0,1]$ and therefore has a maximum, which we denote M_1. Furthermore, $\check{c}(k,z)$ is continuously differentiable. Therefore, $\left|\frac{\partial\check{c}(k,z)}{\partial k}\right|$ is continuous on the compact interval $[\varepsilon,\bar{k}]\times[0,1]$ and is bounded above by some M_2. Define $\bar{\gamma}_k \equiv \sup_{t\in\mathbb{R}}\gamma_k(t)$. Then for any t and $\tilde{k}_1, \tilde{k}_2 \in I$,

$$\begin{aligned}\left|F\left(\tilde{k}_1,t\right) - F\left(\tilde{k}_2,t\right)\right| &\leq \left|f\left(\tilde{k}_1,-t\right) - f\left(\tilde{k}_2,-t\right)\right| + |\delta + \gamma_k(-t)|\cdot\left|\tilde{k}_1 - \tilde{k}_2\right| \\ &\quad + \left|\check{c}\left(\tilde{k}_1,h(-t)\right) - \check{c}\left(\tilde{k}_2,h(-t)\right)\right| \\ &\leq \max\{M_1, M_2, |\delta + \bar{\gamma}_k|\}\left|\tilde{k}_1 - \tilde{k}_2\right|\end{aligned}$$

Taking $\mathcal{K} \equiv \max\{M_1, M_2, |\delta + \bar{\gamma}_k|\}$ completes the proof. □

Lemma 4. *There exists an interval U containing $k_{-\infty}$ in its interior, and a function $w : [a,\infty) \to \mathbb{R}$ such that*

$$\frac{\partial F(k,t)}{\partial k} \geq w(t)$$

for $k \in U$ where $\int_a^\infty w(s)ds > -\infty$.

Proof. Notice that

$$\begin{aligned}\frac{\partial F(k,t)}{\partial k} &= \frac{\partial}{\partial k}\left[-f\left(\tilde{k},-t\right) + (\delta + \gamma_k(-t))\,\tilde{k} + \check{c}\left(\tilde{k},h(-t)\right)\right] \\ &= -\frac{\partial f\left(\tilde{k},-t\right)}{\partial k} + \delta + \gamma_k(-t) + \frac{\partial\check{c}\left(\tilde{k},h(-t)\right)}{\partial\tilde{k}}\end{aligned}$$

If one linearizes $c_-(k)$ around $k_{-\infty}$, one can see from the stable arm that $\frac{\partial c_-(k)}{\partial k}\Big|_{k=k_{-\infty}} = \frac{1}{2}\left(\rho - \gamma_{-u} + \sqrt{(\rho-\gamma_{-u})^2 - 4\frac{c_{-\infty}}{\theta(c_{-\infty})} f''_-(k_{-\infty})}\right)$. Therefore,

$$\begin{aligned}\frac{\partial F(k_{-\infty},\infty)}{\partial k} &= \gamma_{-u} - \rho + \frac{1}{2}\left(\rho - \gamma_{-u} + \sqrt{(\rho-\gamma_{-u})^2 - 4\frac{c_{-\infty}}{\theta(c_{-\infty})} f''(k_{-\infty})}\right) \\ &= \frac{1}{2}\left(\gamma_{-u} - \rho + \sqrt{(\rho-\gamma_{-u})^2 - 4\frac{c_{-\infty}}{\theta(c_{-\infty})} f''_-(k_{-\infty})}\right) > 0\end{aligned}$$

Then by continuity, there exists a $T > 0$ and $\delta > 0$ so that for $k \in (k_{-\infty} - \delta, k_{-\infty} + \delta)$ and $t > T$, $\frac{\partial F(k,t)}{\partial k} > 0$. Therefore, we can take $U = (k_{-\infty} - \delta, k_{-\infty} + \delta)$ and $w(t) = \min\left\{\inf_{k\in U} \frac{\partial F(k,t)}{\partial k}, 0\right\}$. Then $w(t) = 0$ for all $t > T$ so that $\int_a^\infty w(s)ds > -\infty$. □

Proposition 1. *Suppose that Assumption 1 holds, then there exists a unique STraP.*

Proof. We use Theorem 4.7.5 from Hubbard and West (1991) stated in Footnote 1 to show existence and uniqueness of the STraP. Lemma 3 showed that the system is Lipschitz. We need to construct the narrowing upper and lower fence. We restrict attention to a symmetric interval around $k_{-\infty}$ where Lemmas 2 through 4 hold. Define this as $U \equiv [k_{-\infty} - \delta, k_{-\infty} + \delta]$. This is possible as $k_{-\infty}$ was in the interior of all of the intervals.

We describe the explicit construction of the upper fence. The construction of the lower fence proceeds analogously by symmetry. Define $k_0 \equiv k_{-\infty} + \delta$, $k_1 \equiv k_{-\infty} + \frac{\delta}{2}, \cdots,\ k_n \equiv k_{-\infty} + \frac{\delta}{2^n}$ where $n \in \mathbb{Z}_0$. Take $a_n \equiv F_-(k_n)$. Standard growth dynamics tell us that $F_-(k)$ is increasing in k (recall that $F_-(k)$ is the backwards dynamics). Therefore, monotonicity implies that $F_-(k) > a_n$ if $k > k_n$. This also implies that $a_n > 0$ for all n since $F_-(k_{-\infty}) = 0$. By Lemma 3, $F\left(\tilde{k}, t\right)$ gets arbitrarily close to $F_-(k)$. Therefore, for every n, there exists a $T_n > 0$ so that for $t > T_n$,

$$\left|F\left(\tilde{k}, t\right) - F_-\left(\tilde{k}\right)\right| < \frac{a_n}{2}$$

for all $\tilde{k} \in U$.

We can now describe the explicit construction of the upper fence, $\alpha(t)$. Let $\alpha(t)$ be a piecewise linear function starting at time T_1 with $\alpha(T_1) = k_0$ and proceeding linearly to $\alpha(T_2) = k_1$. Then the derivative of α is negative. Furthermore,

$$|F(\alpha(t), t) - F_-(\alpha(t))| < \frac{a_1}{2}$$

and

$$F_-(\alpha(t)) > F_-(k_1) = a_1,$$

for all $t \in [T_1, T_2]$. Therefore, $F(\alpha(t), t) > 0$ for all $t \in [T_1, T_2]$ and $\alpha(\cdot)$ counts as a fence on this interval. Continue concatenating linear functions like this with

$\alpha(T_n) = k_{n-1}$ for all n. $\alpha'(t) < 0$ for all t and $F(\alpha(t), t) > 0$ for all t. Therefore, this piecewise function is a fence converging to $k_{-\infty}$.

Analogously, we can construct a lower fence $\beta(t)$ which converges to $k_{-\infty}$. We have constructed $\alpha(t)$ and $\beta(t)$ that bound an anti-funnel. The final condition is provided by Lemma 4. Thus, we can apply the theorem obtaining the desired result: there is a unique time path that remains in the anti-funnel. This is the STraP. □

B Mapping Special Cases to the General Model

B.1 Baseline Model

As noted in the main body of the paper, the static decisions lead to aggregate Cobb-Douglas production functions, $\mathcal{A}_x(t)K_x(t)^\alpha L_x(t)^{1-\alpha}$ and $\mathcal{A}_c(t)K_c(t)^\alpha L_c(t)^{1-\alpha}$. We therefore focus on the intertemporal equations with these aggregate production functions

and define detrended variables $k(t) \equiv \frac{K(t)}{\mathcal{A}_x(t)^{\frac{1}{1-\alpha}}}$, $x(t) \equiv \frac{X(t)}{\mathcal{A}_x(t)^{\frac{1}{1-\alpha}}}$, $w(t) \equiv \frac{W(t)}{\mathcal{A}_x(t)^{\frac{1}{1-\alpha}}}$ and $c(t) \equiv \frac{C(t)}{\mathcal{A}_c(t)\mathcal{A}_x(t)^{\frac{\alpha}{1-\alpha}}}$. The consumer's problem can then be rewritten

$$\max_{c(t),x(t),k(t),b(t)} \int_{t=\tau}^{\infty} e^{-\rho(t-\tau)} \left(\mathcal{A}_c(t)\mathcal{A}_x(t)^{\frac{\alpha}{1-\alpha}}\right)^{1-\theta} \frac{c(t)^{1-\theta}}{1-\theta} \tag{B.1}$$

s.t.

$$p_c(t)c(t) + p_x(t)x(t) + p_c(t)\dot{b}(t) = w(t)L + R(t)k(t) + r(t)p_c(t)b(t) \tag{B.2}$$

and

$$\dot{k}(t) = x(t) - \left(\delta + \frac{1}{1-\alpha}\frac{\dot{\mathcal{A}}_x(t)}{\mathcal{A}_x(t)}\right)k(t). \tag{B.3}$$

The production side can be simplified to

$$c(t) = L_c(t)^{1-\alpha}k_c(t)^\alpha \tag{B.4}$$

and

$$x(t) = L_x(t)^{1-\alpha}k_x(t)^\alpha. \tag{B.5}$$

It is easy to verify that the First and Second Welfare Theorems both hold so the equilibria correspond with the planner's problem for sufficiently large ρ. The budget constraint is then traded out for the resource constraint

$$c(t) + x(t) = k(t)^\alpha. \tag{B.6}$$

Plugging this into the capital accumulation equation gives

$$\dot{k}(t) = k(t)^\alpha - \left(\delta + \frac{1}{1-\alpha}\frac{\dot{\mathcal{A}}_x(t)}{\mathcal{A}_x(t)}\right)k(t) - c(t) \tag{B.7}$$

The constraint sets are convex and utility is strictly concave so the planner's problem has a unique solution. Finally we note that setting $\gamma_k(t) = \frac{1}{1-\alpha}\frac{\dot{\mathcal{A}}_x(t)}{\mathcal{A}_x(t)}$ and $\mathcal{A}_u(t) \equiv \left(\mathcal{A}_c(t)\mathcal{A}_x(t)^{\frac{\alpha}{1-\alpha}}\right)^{1-\theta}$ fits our framework.

The only thing left is to check the limit conditions and Assumption 1. $f(k,t) = k^\alpha$ and so all limit conditions are automatically satisfied. Notice that if $\gamma_k(t)$ satisfies the limit conditions, then $\gamma_u(t)$ does as well by the algebra of limits and the fact that $\mathcal{A}_c(t)$ is exactly analogous to $\mathcal{A}_x(t)$.

We have that

$$\begin{aligned}\gamma_k(t) &= \frac{1}{1-\alpha}\frac{\dot{\mathcal{A}}_x(t)}{\mathcal{A}_x(t)} \\ &= \frac{1}{1-\alpha}\left[\gamma_x + \frac{\sum_j \omega_{xj} A_j(t)^{\sigma_x-1}\gamma_j}{\sum_j \omega_{xj} A_j(t)^{\sigma_x-1}}\right].\end{aligned} \tag{B.8}$$

With $\sigma_x < 1$, this converges to $\frac{1}{1-\alpha}\left[\gamma_x + \gamma_i\right]$ as $t \to \infty$ where i is the slowest growing sector. Similarly it converges to $\frac{1}{1-\alpha}\left[\gamma_x + \gamma_n\right]$ as $t \to -\infty$ where n is the fastest growing sector. We next check that $\dot{\gamma}_k(t)$ is continuous. Notice that

$$\dot{\gamma}_k(t) = \frac{\sigma_x - 1}{1-\alpha}\left[\frac{\sum_j \omega_{xj} A_j(t)^{\sigma_x-1}\left(\gamma_j\right)^2}{\sum_j \omega_{xj} A_j(t)^{\sigma_x-1}} - \left(\frac{\sum_j \omega_{xj} A_j(t)^{\sigma_x-1}\gamma_j}{\sum_j \omega_{xj} A_j(t)^{\sigma_x-1}}\right)^2\right] \tag{B.9}$$

which is indeed continuous.

We now construct the h function and verify that it satisfies Assumption 1. Notice that $\dot{\gamma}_k(t) < 0$ for all t. Therefore, $\gamma_k : \mathbb{R} \to (\gamma_{-k}, \gamma_{+k})$ is invertible. We can then transform this, $f : \mathbb{R} \to (0,1)$, $f(t) = \frac{\gamma_{-k} - \gamma_k(t)}{\gamma_{-k} - \gamma_{+k}}$. f is then a strictly increasing bijection from $\mathbb{R}$ to $(0,1)$. This will not work as our h function though since the derivative converges to zero at the same rate as $\dot{\gamma}_k(t)$ as $t \to \pm\infty$. For this reason, we will compose it with the function $g : (0,1) \to (0,1)$, $g(x) = (1-x)x^{\frac{1}{2}} + x\left(1 - (1-x)^{\frac{1}{2}}\right)$. This function looks like $x^{\frac{1}{2}}$ as $x \to 0$ and $1 - (1-x)^{\frac{1}{2}}$ as $x \to 1$. Define $h : \mathbb{R} \to (0,1)$ $g(f(t))$.

I can now check the limit conditions using L'Hôpital's rule.

$$\lim_{t\to-\infty} \frac{\dot{\gamma}_k(t)}{h'(t)} = \lim_{t\to-\infty} (\gamma_{-k} - \gamma_k(t))^{\frac{1}{2}}(\gamma_{-k} - \gamma_{+k})^{\frac{1}{2}} = 0,$$

and

$$\lim_{t\to\infty} \frac{\dot{\gamma}_k(t)}{h'(t)} = \lim_{t\to\infty} (\gamma_k(t) - \gamma_{+k})^{-\frac{1}{2}}(\gamma_{-k} - \gamma_{+k})^{\frac{1}{2}} = 0.$$

We also have

$$\lim_{t\to-\infty}\frac{h''(t)}{h'(t)} = \lim_{t\to-\infty}\frac{g'(f(t))f'(t)}{g''(f(t))(f'(t))^2+g'(f(t))f''(t)}$$
$$= \lim_{t\to-\infty}\frac{g'(f(t))}{g''(f(t))f'(t)}+\frac{f'(t)}{f''(t)}$$
$$= \lim_{t\to-\infty}2\frac{\gamma_{-k}-\gamma_k(t)}{\dot\gamma_k(t)}+\frac{\dot\gamma_k(t)}{\ddot\gamma_k(t)}.$$

Notice by L'Hôpital's rule, the limit of $\frac{\gamma_{-k}-\gamma_k(t)}{\dot\gamma_K(t)}$ equals the limit of $-\frac{\dot\gamma_k(t)}{\ddot\gamma_k(t)}$. Then we just find the limit of the first term.

$$\frac{\gamma_{-k}-\gamma_k(t)}{\dot\gamma_k(t)} = \frac{\frac{1}{1-\alpha}\left(\gamma_n-\frac{\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}\gamma_j}{\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}}\right)}{\frac{\sigma_x-1}{1-\alpha}\left[\frac{\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}(\gamma_j)^2}{\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}}-\left(\frac{\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}\gamma_j}{\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}}\right)^2\right]}$$
$$= \frac{\left(\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}(\gamma_n-\gamma_j)\right)\left(\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}\right)}{(\sigma_x-1)\left[\left(\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}(\gamma_j)^2\right)\left(\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}\right)-\left(\sum_j\omega_{xj}A_j(t)^{\sigma_x-1}\gamma_j\right)^2\right]}$$
$$= \frac{\sum_{j,m}\omega_{xj}\omega_{xm}A_j(t)^{\sigma_x-1}A_m(t)^{\sigma_x-1}(\gamma_n-\gamma_j)}{(\sigma_x-1)\sum_{j,m}\omega_{xj}\omega_{xm}A_j(t)^{\sigma_x-1}A_m(t)^{\sigma_x-1}\gamma_j(\gamma_j-\gamma_m)}$$

where γ_n is the fastest growing sector. Then the dominant term is the one with two i's. It converges towards $\frac{1}{(\sigma_x-1)\gamma_n}<0$ as $t\to-\infty$. An exactly analogous argument shows the limit as $t\to\infty$ exists and is positive. Therefore, the baseline model has a STraP.

B.2 Nonhomothetic CES Preferences

This model differs from the baseline model only through the consumption aggregator. $C(t)$ is now implicitly defined by the equation

$$1=\sum_{j=1}^{J}\omega_{cj}^{\frac{1}{\sigma_c}}\left(\frac{C_j(t)}{C(t)^{\varepsilon_j}}\right)^{\frac{\sigma_c-1}{\sigma_c}}. \tag{B.10}$$

The investment sector is the same as before, and so we can detrend capital and investment by the same $\mathcal{A}_x(t)^{\frac{1}{1-\alpha}}$. With these preferences,

$$P_c(t)C(t)+X(t)=\mathcal{A}_x(t)K(t)^{\alpha}L(t)^{1-\alpha}\equiv Y_t. \tag{B.11}$$

Normalizing $L(t)\equiv 1$ for all t, and dividing by $\mathcal{A}_x(t)^{\frac{1}{1-\alpha}}$ gives

$$\frac{P_c(t)C(t)}{\mathcal{A}_x(t)^{\frac{1}{1-\alpha}}}+x(t)=k(t)^{\alpha}. \tag{B.12}$$

Thus we can take $f(k,t) = k^\alpha$. We define $c(t) \equiv \frac{P_c(t)C(t)}{\mathcal{A}_x(t)^{\frac{1}{1-\alpha}}}$.

Then the utility of an agent who chooses $c(t)$ consumption is given by

$$U(c(t),t) = \max_{\{C_j(t)\}} \frac{C(t)^{1-\theta}}{1-\theta}$$

such that

$$1 = \sum_{j=1}^{J} \omega_{cj}^{\frac{1}{\sigma_c}} \left(\frac{C_j(t)}{C(t)^{\varepsilon_j}} \right)^{\frac{\sigma_c - 1}{\sigma_c}}$$

and

$$\sum_j P_j(t) C_j(t) \leq \mathcal{A}_x(t)^{\frac{1}{1-\alpha}} c(t).$$

With the investment sector as the numeraire,

$$P_j(t) = \frac{\mathcal{A}_x(t)}{A_j(t)}.$$

Therefore, the last constraint can be written

$$\sum_j \frac{C_j(t)}{A_j(t)\mathcal{A}_x(t)^{\frac{\alpha}{1-\alpha}}} \leq c(t).$$

This utility function is unbounded and so does not fit into our framework. Thus, we introduce a normalizing factor $N_c(t) \equiv \left(\sum_{j=1}^{J} \omega_{cj} \left(A_i(t) \mathcal{A}_x(t)^{\frac{\alpha}{1-\alpha}} \right)^{\frac{\sigma_c - 1}{\varepsilon_j}} \right)^{\frac{1}{\sigma_c - 1}}$. Then we define $u(c(t),t) \equiv \frac{U(c(t),t)}{N_c(t)^{1-\theta}}$. Therefore,

$$u(c(t),t) \equiv \max_{\{c_j(t)\},\tilde{c}(t)} \frac{\tilde{c}(t)^{1-\theta}}{1-\theta}$$

such that

$$1 = \sum_{j=1}^{J} \omega_{cj}^{\frac{1}{\sigma_c}} \left(\frac{A_j(t)\mathcal{A}_x(t)^{\frac{\alpha}{1-\alpha}}}{N_c(t)^{\varepsilon_j}} \frac{c_j(t)}{\tilde{c}(t)^{\varepsilon_j}} \right)^{\frac{\sigma_c - 1}{\sigma_c}}$$

and

$$\sum_j c_j(t) \leq c(t)$$

where we have replaced $c_j(t) = \frac{C_j(t)}{A_j(t)\mathcal{A}_x(t)^{\frac{\alpha}{1-\alpha}}}$ and $\tilde{c}(t) = \frac{C(t)}{N_c(t)}$.

Therefore $\gamma_k(t)$ is the same as in the baseline model. $\mathcal{A}_u(t) \equiv N_c(t)^{1-\theta}$. This satisfies all of the convergence requirements same as $\mathcal{A}(t)$. As show in Comin et al. (2018), $u(c,t)$ satisfies all of the convergence properties except the ones on the time derivative $\frac{\partial \theta(c,t)}{\partial t} \to 0$. This follows from the fact that it is a differentiable function of the growth rates, the chain rule, and the fact that the derivatives of the growth rates all converge.

B.3 Acemoglu and Guerrieri (2008)

Acemoglu and Guerrieri (2008) differs from the previous two cases by allowing the capital intensity to differ between the sectors. This complication prevents us from aggregating the sectors together. In order to keep the intertemporal problem separate from the intratemporal problem, they assume a single final good that can be used for investment or consumption.

The representative consumer gets utility from consuming the final good

$$\int_{t=-\tau}^{\infty} e^{-\rho(t+\tau)} \frac{C(t)^{1-\theta}}{1-\theta}. \tag{B.13}$$

The single final good is a CES aggregation of multiple sectors

$$Y(t) = \left[\sum_{j}^{J} \chi_j^{\frac{1}{\sigma}} Y_j(t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{B.14}$$

with $\sum_j \chi_j = 1$ and $\chi_j > 0$ for each j. Capital depreciates at a rate $\delta > 0$ so that

$$\dot{K}(t) = Y(t) - \delta K(t) - C(t). \tag{B.15}$$

Each sector produces with a Cobb-Douglas technology

$$Y_j(t) = A_j(t) L_j(t)^{1-\alpha_j} K_j(t)^{\alpha_j}, \tag{B.16}$$

and technological progress is exogenous so that

$$A_j(t) = e^{\gamma_j t} A_j(0) \tag{B.17}$$

with $\gamma_j > 0$ for $j = 1, \dots J$. Capital and labor markets clear so that

$$\sum_{j=1}^{J} K_j(t) = K(t) \tag{B.18}$$

and

$$\sum_{j=1}^{J} L_j(t) = 1. \tag{B.19}$$

As we said, the differing capital intensities implies that we cannot aggregate production. Therefore, we do not have nice expression as we did for $\mathcal{A}_x(t)$ and $\mathcal{A}_c(t)$. However, the ideal detrend function $\mathcal{A}_I(t)$ would solve

$$\mathcal{A}_I(t) = \left[\sum_{j}^{J} \chi_j^{\frac{1}{\sigma}} A_j(t)^{\frac{\sigma-1}{\sigma}} \mathcal{A}_I(t)^{\frac{\alpha_j(\sigma-1)}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{B.20}$$

as this finds how much output actually grows. Using this is not necessary. Instead, we can use the following function which asymptotically agrees with the ideal detrend function, but is easier to use:

$$\mathcal{A}_Y(t) \equiv \left[\sum_j^J \chi_j^{\frac{1}{\sigma}} \left(A_j(t)^{\frac{1}{1-\alpha_j}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}. \tag{B.21}$$

The detrended system is then easy to state. Utility is given by

$$\int_{t=-\tau}^{\infty} e^{-\rho(t+\tau)} \mathcal{A}_Y(t)^{1-\theta} \frac{c(t)^{1-\theta}}{1-\theta}. \tag{B.22}$$

The capital accumulation equation is

$$\dot{k}(t) = y(t) - \left(\delta + \frac{\dot{\mathcal{A}}_Y(t)}{\mathcal{A}_Y(t)}\right) k(t) - c(t). \tag{B.23}$$

Sectoral variables possibly grow at different rates so we detrend production in sector j by $A_j(t)\mathcal{A}_Y(t)^{\alpha_j}$. Then production in sector j is

$$y_j(t) = L_j(t)^{1-\alpha_j} k_j(t)^{\alpha_j}. \tag{B.24}$$

And final production is a time varying function:

$$y(t) = \left[\sum_{j=1}^J \chi_j^{\frac{1}{\sigma}} \left(\frac{A_j(t)}{\mathcal{A}_Y(t)^{1-\alpha_j}} y_j(t)\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}. \tag{B.25}$$

All variables are bounded as $t \to \pm\infty$. Standard arguments then show that the First and Second Welfare theorems both hold.

We now have $\gamma_k(t) = \frac{\dot{\mathcal{A}}_Y(t)}{\mathcal{A}_Y(t)}$ and $\mathcal{A}_u(t) = \mathcal{A}_Y(t)^{1-\theta}$. Just as in the baseline case, we know that $\gamma_k(t)$ monotonically decreases to its limit when $\sigma < 1$. The only difference is it now converges to a capital augmented growth rate. The fastest growing sector is now the sector with the highest $\frac{\gamma_j}{1-\alpha_j}$. As time goes to infinity, γ_k converges to the slowest, and as time goes to $-\infty$ it converges to the fastest. We can also use the same h function as we did in the baseline.

The only complication is verifying that the production function converges. As the first welfare theorem holds, we have that

$$f(k,t) = \max_{L_j(t),k_j(t)} \left[\sum_{j=1}^J \chi_j^{\frac{1}{\sigma}} \left(\frac{A_j(t)}{\mathcal{A}_Y(t)^{1-\alpha_1}} L_j(t)^{1-\alpha_j} k_j(t)^{\alpha_j}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

such that

$$\sum_j L_j(t) = 1$$

and

$$\sum_j k_j(t) = k(t).$$

Denoting the Lagrange multipliers by $\lambda_l(t)$ and $\lambda_k(t)$, the first order conditions are

$$f(k,t)^{\frac{1}{\sigma}}(1-\alpha_j)\chi_j^{\frac{1}{\sigma}}\left(\frac{A_j(t)}{\mathcal{A}_Y(t)^{1-\alpha_i}}L_j(t)^{1-\alpha_j}k_j(t)^{\alpha_j}\right)^{\frac{\sigma-1}{\sigma}}L_j(t)^{-1} = \lambda_l(t)$$

and

$$f(k,t)^{\frac{1}{\sigma}}\alpha_j\chi_j^{\frac{1}{\sigma}}\left(\frac{A_j(t)}{\mathcal{A}_Y(t)^{1-\alpha_j}}L_j(t)^{1-\alpha_j}k_j(t)^{\alpha_j}\right)^{\frac{\sigma-1}{\sigma}}k_j(t)^{-1} = \lambda_k(t).$$

Notice that as $t \to \infty$, $\mathcal{A}_Y(t) \to A_i(t)^{\frac{1}{1-\alpha_i}}$ where i is sector with the slowest augmented growth rate. Therefore, the terms in the sum where $j \neq i$ converge to zero. And on the compact interval $k(t) \in [0, \bar{k}]$, this happens uniformly as we can bound how large those terms are. An analogous argument can be made for $t \, to - \infty$.

We next need to check that the derivative also converges. By the envelope theorem, $\frac{\partial f(k,t)}{\partial k} = \lambda_k(t)$. But then as all activity moves into sector i

$$\begin{aligned}
\lambda_k(t) &= f(k,t)^{\frac{1}{\sigma}}\alpha_j\chi_j^{\frac{1}{\sigma}}\left(\frac{A_j(t)}{\mathcal{A}_Y(t)^{1-\alpha_j}}L_j(t)^{1-\alpha_j}k_j(t)^{\alpha_j}\right)^{\frac{\sigma-1}{\sigma}}k_j(t)^{-1} \\
&\to \left(L_i(t)^{1-\alpha_i}k_i(t)^{\alpha_i}\right)^{\frac{1}{\sigma}}\alpha_i\chi_i^{\frac{1}{\sigma}}\left(L_i(t)^{1-\alpha_i}k_i(t)^{\alpha_i}\right)^{1-\frac{1}{\sigma}}k_i(t)^{-1} \\
&= \alpha_i\chi_i^{\frac{1}{\sigma}}L_i(t)^{1-\alpha_i}k_i(t)^{\alpha_i-1}
\end{aligned}$$

uniformly.

The only thing left to check is that $\frac{\partial f(k,t)}{\partial t}$ and $\frac{\partial^2 f(k,t)}{\partial k \partial t}$ converge to 0 sufficiently fast. This follows from the chain rule. f and $\frac{\partial f}{\partial k}$ are differentiable functions of the productivities. The chain rule then shows that if the derivatives of those growth rates goes to zero, the derivatives of the production functions go to zero uniformly on the compact interval. One can then construct an h function similar to as before.

C Data Construction

The paper utilizes three sets of data, which we discuss in turn.

C.1 U.S. Data

These data are used to yield the calibrations in Sections 4.2 and 4.4, which are in turn used for the simulations in Figures 1-6.

We calibrate the aggregator weights, ω_{cj} and ω_{xj}, $j = a.m.s$, to match the time series average shares of each sector in consumption and investment expenditure. We use the input-output tables to yield the sectoral composition of consumption and investment following closely the procedure described in Herrendorf et al. (2013). To do this, we combined the Make and Use tables for the years 1947-2017 produced by the Bureau of Economic Analysis (BEA) (U.S. Bureau of Economic Analysis, 2016, 2019b).[2] The input-output information from individual sectors is aggregated into our three broad aggregate sectors by simply adding the corresponding row and columns.[3]

To construct the TFP for our three broad sectors, we use data on the real output of each sector, the value added of each sector, and the aggregate capital stock.

The data on the real output of each sector is constructed from the Industry Economic Accounts produced by the BEA (U.S. Bureau of Economic Analysis, 2017, 2019a). In particular, we use data on the value added VA_{lt} and the chain-type price indexes P_{lt}, for disaggregated industries for the years 1947-2018. To do this combines the historical GDP by Industry data for 1947-1997 with the more recent information on the GDP by industry data for 1997-2018. We obtain quantity index for the broad aggregate sectors that we use by aggregating the individual data using the Fisher chain-weighted formula:

$$Q_{j,t} = \left[\frac{\sum_l \frac{P_{lt-1}}{P_{lt}} \text{VA}_{lt}}{\sum_l \text{VA}_{lt-1}} \frac{\sum_l \text{VA}_{lt-1}}{\sum_l \frac{P_{lt}}{P_{lt-1}} \text{VA}_{lt-1}} \right]^{\frac{1}{2}} Q_{j,t-1}.$$

The capital stock series K_t is calculated using the perpetual inventory method. We start with the value of the current-cost net stock of private fixed assets in 1947 as reported by the BEA (U.S. Bureau of Economic Analysis, 2018a), expressed in 2000 prices using the price index for private fixed investment (U.S. Bureau of Economic Analysis, 2019c). To calculate the undepreciated capital we calibrate the depreciation rate δ to match the average depreciation rate of private fixed assets over the period 1947-2017. The depreciation rate of private fixed assets is given by the ratio of the current-cost depreciation of private fixed assets to the current-cost net stock of private fixed assets (U.S. Bureau of Economic Analysis, 2018b). We add to the undepreciated capital stock the gross domestic investment (U.S. Bureau of Economic Analysis, 2018c), expressed in 2000 prices using the price index for private fixed investment.

Given the data on the quantity index for an individual sector Q_{jt}, the value added share share of this sector, $\text{va}_{jt} = \text{VA}_{jt} / \sum_{j'} \text{VA}_{j't}$, and the aggregate capital stock

[2] The Make table refers to the table describing the make of Commodities by Industries, before redefinitions, while the Use table refers to the table given the use of commodities by industries, before redefinitions, using producers' prices.

[3] In particular, we include agriculture, forestry, fishing, and hunting in the broad agriculture sector; mining, utilities, construction, and manufacturing in the broad manufacturing sector; and wholesale trade, retail trade, transportation and warehousing, information, finance, insurance, real estate, rental, and leasing, professional and business services, educational services, health care, and social assistance, arts, entertainment, recreation, accommodation, and food services, other services, except government, and government into the broad service sector.

K_t, the TFP of sector j, inclusive of the contribution of the changes in the aggregate labor input, is given by:

$$A_{jt} = \frac{Q_{jt}}{\text{va}_{jt} \cdot K_t^{\alpha}}.$$

The key assumption to obtain this expression is that the capital can be reallocated without frictions across sectors and that all sectors have the same factor intensities α.

The neutral TFP term affecting the investment aggregator A_{xt} is calibrated using data on the chain-type price indexes of sectoral value added P_{jt} and the Price Index for Private fixed investment P_{xt} using the following expression:

$$A_{xt} = \left[\sum_{j=a,m,s} \omega_{xj} \left(\frac{P_{jt}}{P_{xt}} \right)^{1-\sigma_x} \right]^{\frac{1}{1-\sigma_x}}.$$

Finally, the discount factor is chosen so that the average return to capital in the model between 1950-2000 matches the after tax return to business capital calculated by Gomme et al. (2011).

C.2 Penn World Tables 9.1

The Penn World Tables (PWT) data are used to yield calibration targets for Section 4.5 and for data plotted in Figures 7-9. We detail the sample selection and construction of variables here.

We select the sample based on two criteria. The first criterion is the level of accuracy of the capital stock. Since PWT uses a perpetual inventory method to construct capital, the initial capital is simply assumed and can arbitrarily impact the series during early years as explained in Inklaar et al. (2019). To eliminate this issue, we obtained the additional file "pwt91_services.dta", which contains detailed and disaggregate data on capital and capital services including indicators of the type of series, and included only those series for which the initial capital stock could be constructed from historical data ("t_type"="pim", 38 countries) or those observations sufficiently past the point where initial conditions influence capital stocks ("t_type"="t_star", 92 countries but shorter time series). The second criterion, as noted in the paper, is the level of income because the set of countries thins out considerably as incomes become either very high or very low. We therefore constrain the sample to country-year observations with log real income per capita between 7 (roughly $1100 in 2011 USD) and 10.7464 (roughly $46,500), the U.S. income per capita in 2000. The sample amounts to 126 countries, and 4191 country-year observations.

We plot five different variables in addition to log real income per capita.

- **Real income per capita**: In the model, the population and number of workers are equal, and output and expenditures are also equal. In the real world, however, these differ. Given the importance of nonhomotheticities, we focus on real PPP expenditure income per capita as our measure of development, which we calculate as "rgdpe"/"pop".

- **Capital-output ratio**: Since this is a ratio, we construct these using real data from national accounts as "rnna"/"rgdpna".

- **Investment rate**: We compare this to the current-value investment rate in the model ($p_x X/Y$), so we construct the analog in the data. The PWT report gross capital formation as a share or real output ("cgdpo"), "csh_i", but they do at PPPs rather than current-value shares. We therefore adjust to current value shares using the PPP prices of investment ("pl_i") and output("p_gdpo"). The investment rate is thus calculated as "csh_i"*"pl_i"/"p_gdpo".

- **Relative price of investment**: We construct this using the PPP for investment ("pl_i") and consumption("pl_i") as: "pl_i"/"pl_c".

- **Interest rate**: The PWT 9.1 includes a measure of the internal rate or return, but this does not correspond to our consumption-based interest rate in the model, and there is no simple adjustment that we can make. Instead, we reconstruct the interest rate from scratch following the discrete model's formula for the interest rate in the Euler equation as given in footnote 11 of the paper:

$$\left(\frac{\tilde{C}_{t+1}}{\tilde{C}_t}\right)^{\theta} = \beta\left(1+r_t\right) = \beta\left[1-\delta+\frac{R_{t+1}}{P_{x,t+1}}\right]\left[\frac{P_{x,t+1}/P_{c,t+1}}{P_{x,t}/P_{c,t}}\right]^{1-\theta}.$$

  This requires a depreciation rate, rental rate in units of the capital good, and the growth of the relative price of investment. To back out the rental rate, we start with the detailed disaggregate capital data in the file "pwt91_ services.dta". We follow PWT and construct capital share ("capsh") as one minus labor's share ("labsh") and the share of rent to natural resources ("rntsh"). We then multiply by output ("rgdpna")to get capital payments and divide by the capital stock ("rnna") to get the rental rate *in units of final output*. However, we require the rental rate in units of capital, so we divide by the price of capital relative to output ("pl_n/pl_gdpo").[4] Hence, the full formula for the rental rate is $rentalrate = (1 - \text{"}labsh\text{"} - \text{"}rntsh\text{"}) * \text{"}rgdpna\text{"}/\text{"}rnna\text{"}/(\text{"}pl\_n\text{"}/\text{"}pl\_gdpo\text{"})$. Next, we need to calculate the growth rate in the relative price of investment.

[4] In the model, there is only a single type of investment good so the price of investment and the price of the capital stock are identical. However, the PWT distinguishes between different types of capital, and with a changing composition of investment, the two are distinct, i.e., "pl_n" vs. "pl_i". Equating the returns to investment across capital types, these differences in prices are offset by different depreciation rates. Since our single depreciation rate is used to depreciate the overall capital stock, the price of the overall capital stock is appropriate here. The difference between the two is small, however.

The formula asks for an annual growth rate, but Inklaar et al. (2019) note that these can vary strongly from year to year, so the PWT averages over several years. We therefore follow the PWT in this way and construct it average annual growth rate over q five-year window in the following way

$$relativepricegrowth = (\text{“}pl\_n\text{”}_{t+2}/\text{“}pl\_con\text{”}_{t+2}/(\text{“}pl\_n\text{”}_{t-3}/\text{“}pl\_con\text{”}_{t-3}))^{(1/5)}$$

where we have used t to index the year. Finally, using a depreciation of 0.04, the calibrated depreciation rate in the analysis, we construct the interest rate as

$$r = (1 - 0.04 + rentalrate) * relativepricegrowth - 1.$$

- **10-year real per worker growth**: For growth rates, we utilize real variables in the national accounts, “rgdpna”. Because this is a measure of growth in productivity (rather than living standards), we focus on per worker growth rather than per capita growth. Thus output per worker is “rgdpna”/“emp”. Finally, growth rates are defined both in the model and data in the forward-looking manner, i.e., the growth rate at time t is $(y_{t+10}/y_t)^{1/10} - 1$. Given the forward-looking nature of the variable, we have fewer observations for this constructed variable.

C.3 Groningen Growth and Development Centre 10-Sector Database

These data are used to for the results in Figure 9. We continue our definition of the different sectors: Agriculture (includes Agriculture, forestry, and fishing), Industry (includes Mining and quarrying; Manufacturing; Public utilities; and Construction) and Services (Wholesale and retail trade, hotels and restaurants; Transport, storage, and communication; Finance, insurance, real estate; Community, social and personal services; and Government services). The raw data include 39 countries from 1950-2010, but the value-added data series for some countries are considerably shorter. Current-value shares are constructed by dividing sectoral value added by total value added. Although these data include income and population data from the PWT, the income data are output-side GDP (cgdpo) and from the PWT 8.0 version with the benchmark year of 2005. To keep consistent, we replace these data with the expenditure-side income (cgdpe) from the PWT 9.1, dividing by population (pop) to yield per capita numbers. Once again, we focus on log real incomes per capita between 7 and 10.7464, the U.S. income per capita in 2000. We are left with 36 countries and 1496 country-year observations. The countries are Argentina, Bolivia, Brazil, Botswana, Chile, Colombia, Costa Rica, Denmark, France, Ghana, Hong Kong, Indonesia, India, Italy, Japan, Kenya, Korea, Malawi, Malaysia, Mauritius, Nigeria, Netherlands, Peru, Philippines, Senegal, Singapore, South Africa, Spain, Sweden, Thailand, Taiwan, Tanzania, United States, Venezuala, and Zambia.

D Detailed Derivations for Nonhomothetic CES

In this section, we discuss the steps necessary for the results presented in Section 4.4. We use the discrete version of the model. The production side of the economy is unchanged so that prices, aggregation results, etc. still hold. We thereforce examine the demand side driven by nonhomothetic CES preferences in equation (22). Consider the problem of maximizing C_t subject to the budget constraint $\sum_j P_{j,t}C_{j,t} = E_t \equiv P_{c,t}C_t$. Denoting by λ_t and μ_t the Lagrange multipliers on the definition of preferences and the budget constraint, the first order conditions are

$$\mu_t \omega_j^{\frac{1}{\sigma_c}} \frac{\sigma_c - 1}{\sigma_c} C_{j,t}^{-\frac{1}{\sigma_c}} C_t^{-\epsilon_j \frac{\sigma_c - 1}{\sigma_c}} = \lambda_t P_{j,t}. \tag{D.1}$$

Taking the ratio of (D.1) for two sectors yields (23). Also, multiplying (D.1) by C_t and adding up across all sectors implies that expenditure shares are

$$\frac{P_{j,t}C_{j,t}}{E_t} = \omega_{cj}^{\frac{1}{\sigma_c}} \left[\frac{C_{j,t}}{C_t^{\epsilon_j}}\right]^{\frac{\sigma_c - 1}{\sigma_c}}. \tag{D.2}$$

Next, we consider the dynamic household problem (1) subject to (2) and (3) of Section 2, with the difference that within-period utility is given by the nonhomothetic CES aggregator (22). Combining (2) and (3), we obtain

$$E(C_t) + P_{xt}(K_{t+1} - (1-\delta)K_t) + P_{ct}B_{t+1} = W_t L + R_t K_t + (1 + r_t) P_{ct} B_t \tag{D.3}$$

where

$$E_t(C_t) = \left[\sum_j \omega_j \left(P_{j,t} C_t^{\epsilon_j}\right)^{1-\sigma_c}\right]^{\frac{1}{1-\sigma_c}} \tag{D.4}$$

denotes the total expenditure in consumption. Noting that[5]

$$\frac{\partial E_t(C_t)}{\partial C_t} = \bar{\epsilon} P_{c,t} \tag{D.5}$$

where $\bar{\epsilon}_t = \sum_j \epsilon_j \frac{P_{j,t}C_{j,t}}{E_t}$, the Euler equation becomes

$$\left(\frac{C_{t+1}}{C_t}\right)^\theta = \beta \frac{\bar{\epsilon}_t}{\bar{\epsilon}_{t+1}} \frac{P_{c,t}}{P_{c,t+1}} \frac{R_{t+1} + P_{x,t+1}(1-\delta)}{P_{x,t}} = \beta \frac{\bar{\epsilon}_t}{\bar{\epsilon}_{t+1}} \left(1 - \delta + \frac{R_{t+1}}{P_{x,t+1}}\right) \frac{\frac{P_{c,t}}{P_{x,t}}}{\frac{P_{c,t+1}}{P_{x,t+1}}}. \tag{D.6}$$

[5]This result comes from the observation that $\frac{\partial E_t}{\partial C_t} = \frac{\partial P_{c,t}(C_t)}{\partial C_t} C_t + P_{c,t}$ and

$$\frac{\partial P_{c,t}(C_t)}{\partial C_t} C_t = P_{c,t}^{\sigma_c} \sum_j \omega_j (\epsilon_j - 1) \left(P_{j,t} C_t^{\epsilon_j - 1}\right)^{1-\sigma_c} = P_{c,t}^{\sigma_c - 1 + 1} C_t^{\sigma_c - 1} \sum_j \omega_j (\epsilon_j - 1) \left(P_{j,t} C_t^{\epsilon_j}\right)^{1-\sigma_c} = P_{c,t}(\bar{\epsilon} - 1).$$

Disucssion on the asymptotic behavior Our analysis maintains the assumptions that $\epsilon_a < \epsilon_m < \epsilon_s$ and that $\gamma_a > \gamma_m > \gamma_s$. On the investment side, we already argued that this implies that, as $t \to \infty(-\infty)$, the investment share of agriculture (services) tends to one. For expenditure shares, we also noted that our preference specification

$$1 = \sum_{j=a,m,s} \omega_{cj}^{\frac{1}{\sigma_c}} \left[\frac{C_{j,t}}{C_t^{\epsilon_j}} \right]^{\frac{\sigma_c - 1}{\sigma_c}} = \sum_j x_{jt},$$

where x_{jt} denotes expenditure shares. Next, we use the same reasoning as in Comin et al. (2015) to derive asymptotic growth rates. We have that

$$\frac{P_{j,t}C_{j,t}}{P_{i,t}C_{i,t}} = \frac{\omega_j}{\omega_i} \left(\frac{P_{j,t}}{P_{i,t}} \right)^{1-\sigma_c} C_t^{(\epsilon_j - \epsilon_i)(1-\sigma_c)} = \frac{\omega_j}{\omega_i} \left(\frac{A_{i,t}}{A_{j,t}} \right)^{1-\sigma_c} C_t^{(\epsilon_j - \epsilon_i)(1-\sigma_c)} \tag{D.7}$$

Consider the case that $t \to \infty$. The term corresponding to prices tends to zero as long as $j = a, m$ and $i = s$. Also, as utility is increasing unboundedly in each elements of $\{C_{it}\}_{i=a,m,s}$, we have that $C_t \to \infty$. This implies that $C_t^{(\epsilon_j - \epsilon_i)(1-\sigma_c)} \to 0$ if $j = a, m$ and $i = s$. Thus, price and income effects work in the same direction towards making relative expenditure shares being 0 for $j = a, m$. Using the fact that expenditure shares add up to one, this implies that the expenditure share in services tends asymptotically to 1. An analogous argument for $t \to -\infty$ shows that in that limit expenditure share in agriculture tends to 1 and it is zero for manufacturing and services. In sum, the asymptotic properties of the dynamical system remain the same as in our baseline exercise.

Using that asymptotically the surviving sector accounts for the entire economy, we can derive the asymptotic growth rates of the consumption aggregator from the expenditure share definition and its growth rate,

$$1 = \frac{\left(\frac{P_{j,t+1}}{E_{t+1}} \right)^{1-\sigma_c} C_{t+1}^{(1-\sigma_c)\epsilon_j}}{\left(\frac{P_{j,t}}{E_t} \right)^{1-\sigma_c} C_t^{(1-\sigma_c)\epsilon_j}} \tag{D.8}$$

which implies that

$$\frac{C_{t+1}}{C_t} = \left[\frac{\left(\frac{E_{t+1}}{E_t} \right)}{\left(\frac{P_{j,t+1}}{P_{j,t}} \right)} \right]^{1/\epsilon_j} = \left[\frac{((1+\gamma_x)(1+\gamma_j))^{1/(1-\alpha)}}{(1+\gamma_x)} \right]^{1/\epsilon_j} = \left[(1+\gamma_x)^{\alpha/(1-\alpha)} (1+\gamma_j)^{1/(1-\alpha)} \right]^{1/\epsilon_j}. \tag{D.9}$$

Asymptotically, we also have that $\bar{\epsilon}_t = \bar{\epsilon}_{t+1}$. From this observation, it follows that the Euler equation satisfies,

$$(1+\gamma_C)^{\theta-1} = \beta \left(1 - \delta + R \right) \left((1+\gamma_x)(1+\gamma_j) \right)^{-1/(1-\alpha)} \tag{D.10}$$

where R is the asymptotic level of interest rate, γ_C is the asymptotic growth rate of real consumption and j is the surviving sector. Combining the definition of gross interest rate,

$$R = \alpha \tilde{K}^{\alpha-1} \tag{D.11}$$

where $\tilde{K} = K_t / \mathcal{A}_{x,t}^{1/(1-\alpha)}$ with the Euler Equation (24), we obtain that

$$\tilde{K} = \left(\frac{\alpha\beta}{(1+\gamma_x)^{\frac{\alpha(\theta-1)+\epsilon_j}{\epsilon_j(1-\alpha)}} (1+\gamma_j)^{\frac{\theta-1+\epsilon_j}{\epsilon_j(1-\alpha)}} - \beta(1-\delta)} \right)^{1/(1-\alpha)} \tag{D.12}$$

where j denotes the growth of the sector surviving asymptotically. For $\epsilon_j = 1$, this expression becomes our baseline result,

$$\tilde{K} = \left(\frac{\alpha\beta}{(1+\gamma_x)^{\frac{\alpha\theta+1-\alpha}{(1-\alpha)}} (1+\gamma_j)^{\frac{\theta}{(1-\alpha)}} - \beta(1-\delta)} \right)^{1/(1-\alpha)} . \tag{D.13}$$

Pseudo-Balanced Growth Path We define the pseudo-balanced growth path as the balanced growth path to which the economy would converge if the present level of growth was maintained going forward, implying that normalized expenditures in the pseudo-BGP are

$$\tilde{E}_0 = \tilde{K}_0^\alpha - \tilde{K}_0 \left((1+\gamma_{Ax})^{1/(1-\alpha)} + \delta - 1 \right) . \tag{D.14}$$

We proceed by assuming that nominal variables grow at the rate implied by $\mathcal{A}_{x,0}$ and infer the growth rate of real consumption at time $t+1$ so that it is consistent with the "pseudo-BGP" capital level,

$$\left(\frac{C_{t+1}}{C_t} \right)^{\theta-1} = \beta \frac{\bar{\epsilon}_t}{\bar{\epsilon}_{t+1}} \left(1 - \delta + \alpha \tilde{K}_0^{\alpha-1} \right) (1+\gamma_{Axt})^{-1/(1-\alpha)}. \tag{D.15}$$

In contrast to the homothetic case, since we do not know the growth rate of $P_{c,t}$, we cannot apply the formula for the pseudo-BGP capital level, Equation (D.12). Instead, we jointly solve for the Euler equation and the expenditure functions at time t and $t+1$ to pin down $\bar{K}_0$, C_t and C_{t+1}. Substituting the definition of the pseudo-BGP into the Euler Equation in Footnote 19 yields

$$1 = (1+\gamma_{Ax})^{\frac{-\sigma_c}{1-\alpha}} \beta \left(1 - \delta + \alpha \tilde{K}_0^{\alpha-1} \right) \left(\frac{C_t}{C_{t+1}} \right)^{\theta-1} \frac{\sum_j \omega_{cj} \left(P_{j,t}/P_{x,t} \right)^{1-\sigma_c} C_t^{(1-\sigma_c)\epsilon_j} \epsilon_j}{\sum_j \omega_{cj} \left(P_{j,t+1}/P_{x,t+1} \right)^{1-\sigma_c} C_{t+1}^{(1-\sigma_c)\epsilon_j} \epsilon_j}. \tag{D.16}$$

We use this expression with the definition of the expenditure function at t and $t+1$ combined with the expression for expenditures in the "pseudo-BGP", Equation

(D.14). For time t, we have

$$\left[\sum_j \omega_j \left(P_{j,t} C_t^{\epsilon_j}\right)^{1-\sigma_c}\right]^{\frac{1}{1-\sigma_c}} = \mathcal{A}_t^{1/(1-\alpha)} \left(\tilde{K}_0^\alpha - \tilde{K}_0 \left((1+\gamma_{Ax})^{1/(1-\alpha)} + \delta - 1\right)\right). \tag{D.17}$$

Equation (D.17) evaluated at t and $t+1$ with the Euler Equation (D.16) constitute a system of three equations in three unknowns, $\tilde{K}_0$, C_t, and C_{t+1}. We solve numerically for these values.

Some practical computational details For both the pseudo-BGP and the shooting algorithm, it is convenient to normalize C_t. This allows to numerically compute C_t for any t while keeping the range of potential values for the solver in a bounded interval. Note that C_t asymptotes to a constant growth rate as $t \to \pm\infty$ that is different from the asymptotic growth of nominal values. We thus introduce the normalizing factor that combines the asymptotic growth rates of each sector combined as a Holder-mean,

$$\begin{aligned} N_{c,t} &= \left(\sum_{i=a,m,s} \omega_c \left[(1+\gamma_x)^{\alpha/(1-\alpha)} (1+\gamma_i)^{1/(1-\alpha)}\right]^{-\frac{t}{\epsilon_i}(1-\sigma_c)}\right)^{-1/(1-\sigma_c)} \\ &= \left(\sum_{i=a,m,s} \omega_c \left[{A_{x,t}}^\alpha A_{i,t}\right]^{\frac{(\sigma_c-1)}{\epsilon_i(1-\alpha)}}\right)^{-1/(1-\sigma_c)}, \end{aligned} \tag{D.18}$$

so that it asymptotes to the same growth rates as C_t. As a result, it is possible to normalize consumption by this factor, $\tilde{C}_t = C_t/N_{c,t}$ so that it converges to a constant number as $t \to \pm\infty$. This allows us to define the asymptotic values of normalized consumption, $\tilde{C}_{-\infty}$ and $\tilde{C}_\infty$. They provide a natural scale for the search range of any solver algorithm.

We illustrate this by writing the the forward Euler Equation we use to solve for C_{t+1} in the shooting algorithm,

$$\begin{aligned} &\beta\left(1-\delta+R_{t+1}\right)\left(\tilde{E}_t\right)^{\sigma_c}\left(\frac{\mathcal{A}_{t+1}^{1/(1-\alpha)}}{\mathcal{A}_t^{1/(1-\alpha)}}\right)^{-\sigma_c} C_t^{\theta-1}\sum_j \omega_{cj}\left(P_{j,t}\right)^{1-\sigma_c} C_t^{(1-\sigma_c)\epsilon_j}\epsilon_j = \\ &\tilde{E}_{c,t+1}^{\sigma_c}\left(C_{t+1}\right)^{\theta-1}\sum_j \omega_{cj}\left(P_{j,t+1}\right)^{1-\sigma_c} C_{t+1}^{(1-\sigma_c)\epsilon_j}\epsilon_j. \end{aligned} \tag{D.19}$$

D.1 Estimation of the Nonhomothetic CES parameters

We use the intra-period allocation of consumption across sector to estimate the price and income elasticity parameters (up to a normalization constant). We rewrite the

Table D.1: Estimated Parameters NHCES, US Consumption VA 1947-2017

	Unconstrained	Imposing $\sigma = 0.01$
σ	0.30 (0.09)	–
ϵ_a	0.43 (0.22)	0.92 (0.04)
ϵ_s	1.64 (0.17)	1.30 (0.01)
$\ln \omega_a$	2.39 (1.36)	-1.90 (0.54)
$\ln \omega_s$	-4.57 (0.78)	-2.74 (0.17)

Estimation with the normalization $\epsilon_m = \omega_m = 1$.

demand system in terms of observables,

$$\ln\left(\frac{P_{it}C_{it}}{\sum_j P_{jt}C_{jt}}\right) = \ln\left(\frac{\omega_i}{\omega_l^{\frac{\epsilon_i}{\epsilon_l}}}\right)(1-\sigma)\ln\left(\frac{P_{it}}{P_{lt}}\right) + (1-\sigma)\left(\frac{\epsilon_i}{\epsilon_l}-1\right)\ln\left(\frac{\sum_j P_{jt}C_{jt}}{P_{lt}}\right) + \frac{\epsilon_i}{\epsilon_l}\ln\left(\frac{P_{lt}C_{lt}}{\sum_j P_{jt}C_{jt}}\right)$$

for all $i \neq l$. We normalize $\epsilon_m = \omega_m = 1$ and estimate $\{\sigma, \epsilon_a, \epsilon_s, \omega_a, \omega_s\}$. Since a priori any sector can be used as reference l we stack the demand equations using different sectors as reference (as in Comin et al., 2019) and find the parameters that generate the best fit of the data using a GMM procedure.[6] The resulting estimated elasticity parameters are $\sigma = 0.30$, $\epsilon_a = 0.43$ and $\epsilon_s = 1.64$ as shown in Table D.1.

References

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

Acemoglu, D. and V. Guerrieri (2008): "Capital Deepening and Nonbalanced Economic Growth," *Journal of Political Economy*, 116, 467–498.

Comin, D., A. Danieli, and M. Mestieri (2019): "Income-Driven Labor-Market Polarization," 2019 Meeting Papers 1398, Society for Economic Dynamics.

Comin, D. A., D. Lashkari, and M. Mestieri (2015): "Structural Change with Long-run Income and Price Effects," Working Paper 21595, National Bureau of Economic Research.

[6]We use manufacturing and services as reference sectors and stack the equations that result from using these sectors as references in our estimation. We exclude agriculture as it represents a very small share of value added consumption and having it as a reference introduces noise in the estimation, resulting in a worse residual sum of squares relative to only using manufacturing and services as references. We also report the estimation results of ϵ's fixing $\sigma = 0.01$.

——— (2018): "Structural Change with Long-run Income and Price Effects," Working Paper 21595, National Bureau of Economic Research.

Gomme, P., B. Ravikumar, and P. Rupert (2011): "The Return to Capital and the Business Cycle," *Review of Economic Dynamics*, 14, 26–278.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2013): "Two Perspectives on Preferences and Structural Transformation," *American Economic Review*, 103, 2752–89.

Hubbard, J. and B. West (1991): *Differential equations: a dynamical systems approach. Part I: ordinary differential equations*, Springer-Verlag.

Inklaar, R., P. Woltjer, and D. Gallardo Albarrán (2019): "The Composition of Capital and Cross-country Productivity Comparisons," *International Productivity Monitor*, 34–52.

U.S. Bureau of Economic Analysis (2016): "Historical Make-Use Tables, 1947-1996," https://www.bea.gov/industry/input-output-accounts-data, release Date: February 19, 2016.

——— (2017): "GDP by Industry, Historical Data, 1947-1997," https://www.bea.gov/data/gdp/gdp-industry, release Date: March 03, 2017.

——— (2018a): "Table 1.1. Current-Cost Net Stock of Fixed Assets and Consumer Durable Goods," https://www.bea.gov/data/investment-fixed-assets/by-type, release Date: November 20, 2018.

——— (2018b): "Table 1.3. Current-Cost Depreciation of Fixed Assets and Consumer Durable Goods," https://www.bea.gov/data/gdp/gross-domestic-product, release Date: November 20, 2018.

——— (2018c): "Table 5.2.5. Gross and Net Domestic Investment by Major Type," https://www.bea.gov/data/gdp/gross-domestic-product, release Date: July 31, 2018.

——— (2019a): "GDP by Industry, 1997-2018," https://www.bea.gov/data/gdp/gdp-industry, release Date: April 19, 2019.

——— (2019b): "Make-Use Tables, Summary Level, 1997-2017," https://www.bea.gov/industry/input-output-accounts-data, accessed 26 September 2019.

——— (2019c): "Table 5.3.4. Price Indexes for Private Fixed Investment by Type," https://www.bea.gov/data/gdp/gross-domestic-product, release Date: April 26, 2019.