From Smith to Schumpeter: A Theory of Take-off and Convergence to Sustained Growth

Pietro Peretto

Duke University

June 4, 2013

ERID Working Paper Number 148

This paper can be downloaded without charge from the Social
Science Research Network Electronic Paper Collection:

<http://ssrn.com/abstract=2275363>

**Economic Research Initiatives at Duke
WORKING PAPERS SERIES**

From Smith to Schumpeter: A Theory of Take-off and Convergence to Sustained Growth

Pietro F. Peretto*

Department of Economics, Duke University

June 4, 2013

Abstract

This paper develops a theory of the emergence of modern innovation-driven Schumpeterian growth. It uses a tractable model that yields a closed-form solution, consisting of an S-shaped (i.e., logistic-like) time path of firm size and a set of equations that express the relevant endogenous variables — GDP, product variety and product quality, consumption, the shares of GDP earned by the factors of production — as functions of firm size. It also obtains closed-form solutions for the dates of the events that drive the economy's phase transitions as functions of the fundamentals. The resulting path of GDP per capita consists of a convex-concave profile replicating the key feature of long-run data: an accelerating phase followed by a deceleration with convergence to a stationary growth rate. Compared to other available theories, the paper focuses on the *within-industry* forces that regulate the response of firms and entrepreneurs to Smithian market expansion.

Keywords: Endogenous Growth, Firm Size, Market Structure, Take-off.

JEL Classification Numbers: E10, L16, O31, O40

*Address: Department of Economics, Duke University, Durham, NC 27708. Phone: (919) 6601807. Fax: (919) 6848974. E-mail: peretto@econ.duke.edu. I thank for comments and suggestions: John Seater, Domenico Ferraro, David Weil, Oded Galor, Philippe Aghion, Ufuk Akcigit, Paul Romer, Michael Peters and, especially, Peter Howitt.

Technological creativity seems to be a uniform and ubiquitous feature of the human species, and yet just once in history has it led to a sea change comparable to phase transitions in physics or the rise of *Homo sapiens sapiens* in evolutionary biology. The Industrial Revolution and the subsequent developments did not just rise the *level* of technological capabilities; they changed the entire dynamics of how innovation comes about and the speeds of both invention and diffusion. For much of human history, innovation had been primarily a byproduct of normal economic activity, punctuated by periodical flashing insight that produced a macroinvention, such as water mills or the printing press. But sustained and continuous innovation resulting from systematic R&D carried out by professional experts was simply unheard of until the Industrial Revolution. (Mokyr 2010, p. 37)

1 Introduction

One of the most interesting questions tackled by contemporary research on economic development and growth concerns the forces that drove the massive acceleration of income per capita growth at the time of the Industrial Revolution. The literature has produced insights that have changed the way we think about the issue.¹ Current theories, however, do not yet rise to what Mokyr (2005, 2010), among others, sees as the main challenge: to explain not only the rise of the growth rate, but also the qualitative transformation of the economy as sustained growth fueled by profit-driven technological change become its defining feature. To quote in full:

“Pre-1750 growth was primarily based on Smithian and Northian effects: gains from trade and more efficient allocations due to institutional changes. The Industrial Revolution, then, can be regarded not as the beginnings of growth altogether but as the time at which technology began to assume an ever-increasing weight in the generation of growth and when economic growth accelerated dramatically” (2005, p. 1118).

In this paper I develop a theory of the emergence of modern growth. The main building block is a Schumpeterian model with two types of innovation activity: existing firms invest in-house to improve the quality of the goods they sell; entrepreneurs invest to design new products and set up new firms to serve the market.² I refer to these activities as quality (or vertical) and variety (or horizontal) innovation, respectively. The literature has long recognized that the returns to these

¹In the last ten, fifteen years this literature has flourished and is now immense. For reviews, see Galor (2005, 2011). Mokyr and Voth (2010) discuss its contribution to improving the historian’s understanding of the issues.

²Why a Schumpeterian model? In Mokyr’s words, history provides *the* answer (emphasis mine): “...favorable institutions explain first and foremost the kind of Smithian growth in which the expansion of commerce, credit, and more labor mobility were the main propulsive forces. But the exact connection between institutional change and the rate of innovation seems worth exploring, precisely because the Industrial Revolution marked the end of the old regime in which economic expansion was driven by commerce and *the beginning of a new Schumpeterian*

activities are increasing in the scale of operation of the firm. What it has not fully appreciated is that because investment in innovation is a sunk cost that is economically justified only when the anticipated revenue flow is sufficiently large, there exist corner solutions where investment in innovation is zero.³ Taking this property into account helps in meeting Mokyr’s challenge because it produces a theory of *when* and *how* endogenous Schumpeterian innovation becomes the main driver of growth.

To appreciate the paper’s contribution, it is instructive to compare its perspective to the Unified Growth Theory (UGT) pioneered by Galor and Weil (2000) and reviewed in Galor (2005, 2011; see also Lucas 2002), which currently is the dominant approach to the study of the Industrial Revolution.

First, the focus is different. Rather than studying the breakdown of the Malthusian regime and the demographic transition, this paper focuses on the incentives to innovate and how they evolve with aggregate market size. This difference in focus drives its different modeling choices. Specifically, in order to focus on the role of the quantity-quality trade-off for children, UGT needs to simplify elsewhere — and it does so by modeling technological change as a black box. This paper does the opposite: it opens the black box of technological change and keeps the model tractable by abstracting from reproduction and education decisions.⁴

Second, the core logic of the Malthusian regime is that population growth is the channel through which technological change yields larger output while living standards stagnate. The idea is that population size fully absorbs improvements in the economy’s production capacity. Crucially, therefore, the Malthusian regime says that the economy develops and eventually activates the quantity-quality trade-off for children only *because* there is technological change, either exogenous or associated to population size, that drives the effective supply of land.⁵ But then, precisely because it

world of innovation” (2010, pp. 37-38). Another, equally important, reason is that this class of models has received substantial empirical support recently, especially as an explanation of long-term historical data. For examples, see Ha and Howitt (2007), Madsen (2008, 2010), Madsen and Ang (2011), Madsen, Ang and Banerjee (2011), Laincz and Peretto (2006), Ulku (2007).

³Such corner solutions exist in all models of endogenous innovation but to date have played no role in the theoretical work on the Industrial Revolution.

⁴Recent notable attempts to integrate endogenous innovation mechanisms in UGT are Desmet and Parente (2012) and Strulik, Prettnner and Prskawetz (2013). Although they ask different research questions, both papers propose full-fledged UGT models that, because of their ambition and large number of key ingredients, are quite complex. Desmet and Parente (2012) ask a question and deploy a mechanism that are sufficiently close to what this paper does to deserve some discussion. They argue that a central mechanism at the heart of the Industrial Revolution was the rise in market competition, visible in falling mark-ups. They credit Peretto (1998, 1999a, 1999b) for being the first to integrate such mechanisms in endogenous growth theory, and extend the scope of the analysis by adding the key ingredients of UGT (e.g., fertility choice) and an agricultural sector whose productivity grows exogenously in the early history of the model. Given the complexity of the model, they need to resort to numerical simulations. This paper, in contrast, shuts down those additional UGT channels and concentrates on the industry-level Schumpeterian response to market expansion and provides an analytically transparent characterization of the forces at play. Moreover, because it shuts down endogenous mark-ups as well, it emphasizes different drivers of the incentives to innovation.

⁵UGT models this process in reduced-form by positing that the rate of land-augmenting technological change is an increasing function of population size that yields a strictly positive rate of technological change for all values of

models technological change as a black box, UGT leaves unexplained the key driver of the dynamics. To see why this is important, note that if, applying the Malthusian logic, the economy studied in this paper sets population growth at zero to stabilize output per capita, it cannot take off since it fails to cross the threshold of market size for positive technological progress. This observation forces out in the open the danger of the back-box assumption in UGT and, perhaps more interestingly, it suggests that initially shrinking output per capita due to a growing population is the price that society pays to create the aggregate market size needed to support profit-driven investment in innovation. This feature seems to turn the logic of Malthusian equilibria on its head and is worth exploring since it refines our understanding of the dynamic interactions among land scarcity, technology and demography.

A third difference between UGT and this paper, therefore, is that this paper posits that population growth drives aggregate market growth and eventually takes the economy across the threshold of *firm size* — which is really a threshold of profitability — where investment in new technology yields a sufficiently high rate of return. In this perspective, the paper articulates a vision of the take-off process in line with that proposed by Julian Simon (2000), who argued that (exogenous) population growth triggered the “great breakthrough” and the consequent acceleration of world growth. The paper’s characterization, which yields an analytical solution for the growth path, sheds new light on long-standing questions concerning the features of the process. Why is this difference important? Aside from it being a modeling simplification analogous to UGT’s black boxing of technological change — and thus something to be explicitly acknowledged — this assumption cuts to the chase and focuses directly on the role of population growth as the trigger of the economy’s phase transition independently of the underlying forces driving it. Furthermore, it fully acknowledges that the existence of a corner solution where technological change is zero suggests the potential for a chicken-and-egg problem: which comes first, technological change or population growth? UGT takes the view that land-augmenting technological change comes first and drives (i.e., causes) population growth; this paper simplifies things by going to the other extreme: population expansion triggers the onset of Schumpeterian innovation.

To summarize, the key features of my analysis are the following. First, to keep thing simple the model takes no position on the initial state of the economy: it simply posits population growth and studies its implications.⁶ Moreover, it posits that an activity already recognizable as “industry”

population size. If, instead, it admits a threshold level of population below which technological change is zero, then escaping the Malthusian trap is not a necessary outcome of the model. In fact, such a variant of the theory would yield the same prediction as the one developed here. Namely, the economy needs an initial period of population growth *not* driven by technological change to cross the population size threshold and activate the engine of growth.

⁶One interpretation is that population growth just stands in for forces that enlarge the market. Alternatives to exogenous population growth that I have explored are: exogenous disembodied technological change; exogenous growth of the resource endowment (e.g., discovery and opening up of new land); growth of embodied knowledge through “natural” curiosity and/or learning by being/doing effects. It is possible, moreover, to augment the model with endogenous fertility and reproduce the main results discussed here. An advantage of such exercise is that it captures additional feedbacks that I leave out of this paper for simplicity. They are nevertheless worth studying, and I am doing so in work in progress. It is also possible to add features that make the initial state Malthusian in the

undertaken in institutions already recognizable as “firms” exists.⁷ These are simplifications that allow me to concentrate on the forces that drive firms’ and entrepreneurs’ incentives to engage in costly innovation activity. Second, the model uses population growth as the trigger that moves the economy from a state of affairs with no profit-driven innovation to one with it. The Schumpeterian approach, however, shows that it is not market size per se that matters, but its contribution to firm size through interactions that thus far have been ignored. In this perspective, the paper focuses on the *within-industry* forces that regulate the response of firms and entrepreneurs to Smithian market expansion and identifies an amplification mechanism that is not specific to a particular driver — population growth or something else — of such expansion. The paper thus differs from UGT fundamentally in that it puts *firms* and *industry* at the heart of the theory of the Industrial Revolution that it proposes.

The paper is organized as follows. Section 2 provides a preview of the main results. Section 3 sets up the model. Section 4 solves it. Section 5 interprets the solution and discusses its implications for history and its potential empirical applications. Section 6 discusses further features of the theory that although not central to the paper’s main message are nevertheless of interest. Section 7 concludes.

2 Preview of the results

As mentioned in the previous section, I posit exogenous population growth to capture, in reduced form, forces that enlarge the market in the economy’s early history. This is a simplification that proves convenient in deriving analytical results and in focusing the paper on the evolution of industrial activity. There exist two thresholds of firm size, one where variety innovation is zero and the other where quality innovation is zero. Because at least one of these two thresholds is finite, as long as the market for industrial goods grows due to aggregate forces, the economy *must* turn on Schumpeterian innovation. The intuition is that the rents earned by incumbent firms become larger and larger and eventually must be competed away either by entry of new firms or by in-house investment by existing firms.

Since innovation must start, the only question is *when* and what specific *sequence of events* unfolds. The model reduces to a pair of piece-wise linear differential equations describing the evolution of firm size in two scenarios. In one the economy crosses the threshold for variety innovation first, in the other it crosses the threshold for quality innovation first. In each scenario I obtain a closed-form solution, consisting of an S-shaped (i.e., logistic-like) time path of firm size and a set of equations that express the relevant endogenous variables — GDP, product variety and product quality, consumption, the shares of GDP earned by the factors of production — as functions of firm size. I also obtain closed-form solutions for the dates of the events that drive the economy’s phase

sense of UGT, but, for the purposes of this paper, doing so complicates the analysis with no gain in insight. I provide further details on the comparison between this model and UGT in the Conclusion section.

⁷It is possible to add features that explain the appearance of manufacturing firms undertaking production in centralized facilities called factories, but doing so would take the analysis beyond its present scope.

transitions as functions of the fundamentals. The transition path of GDP per capita consists of a convex-concave profile replicating the key feature of long-run data: an accelerating phase followed by a deceleration with convergence from below to a stationary growth rate.

The story that these solutions tell is one where the economy starts out in a situation where there is no profit-driven innovation and firms earn rents that grow with the size of the market. They also reap efficiency gains due to static economies of scale (i.e., unit production costs fall with the volume of production). There is no guarantee, however, that such gains translate aggregate growth into per capita growth. Moreover, whatever its sign, the growth rate of GDP per capita in this phase is negligible since it is a fraction of the rate of population growth.

What happens next depends on which type of innovation starts first. If variety innovation starts first, there is a period in which the tension between the exploitation of firm-level static economies of scale, that requires firm growth, and the exploitation of social returns to variety, that requires entry, results in a profile of GDP per capita growth that is always convex but can be increasing, U-shaped or decreasing over time. Hence, the onset of profit-driven horizontal innovation can be, but not necessarily is, associated to a continuation of the slowdown due to the gradual exhaustion of static economies of scale. This intermediate phase ends when the economy crosses the threshold for quality innovation. The solution for the date of this event says that it is not necessarily finite so that the economy may fail to complete the transition to modern growth.

If, instead, quality innovation starts first, there is a period in which the rate of innovation exhibits explosive behavior because firms are still earning escalating rents driven by aggregate market growth. This intermediate phase has *finite duration* because the date when the economy crosses the threshold for variety innovation is necessarily finite. The time profile of the growth rate of GDP per capita is necessarily convex and decreasing at the onset of quality innovation. The reason is that the initial contribution of quality growth cannot overcome the gradual exhaustion of Smithian static economies of scale since initially it follows a very shallow path.

In both cases growth eventually accelerates, as the contribution of Schumpeterian innovation starts dominating over the gradually vanishing contribution of Smithian economies of scale. Modern growth takes hold when firm size is sufficiently large and the economy turns on both innovation engines. In this final phase the economy exhibits desirable properties, like the sterilization of the scale effect, that have interesting implications for the role of fundamentals and policy.

The closed-form solution for the transition path provides analytical insight on the timing of the key events in the economy's history. It identifies the determinants of the take-off date, defined as the onset of profit-driven innovation, and the determinants of the duration of the transition to the modern phase with both innovation engines turned on and convergence to sustained scale-free growth. Moreover, it provides a novel insight on why the economy might fail to reach the modern growth phase: when horizontal innovation starts first, it might fail to turn on vertical innovation due to premature market saturation.

The conflicting forces playing out in the intermediate phase result in a rich set of possible shapes of a path that eventually converges to the general S-shaped pattern described above. The

comparison of these possible histories suggests that the cross-country variation that the data show in terms of take-off dates and shapes of the paths in the immediate neighborhood of the take-off date should not come as a surprise and can be explained within a unified analytical structure that captures the interplay between the size of the market, which evolves according to aggregate forces like population, and the profitability of the individual firm that, given market size, depends on the number of firms active in the market.

3 The model

The economy is closed. All variables are functions of (continuous) time but to simplify the notation I omit the time argument unless necessary to avoid confusion.

3.1 Households

The economy is populated by a representative household that supplies labor and land services and trades assets in competitive markets. The household has preferences

$$U(0) = \int_0^\infty e^{-\rho t} L(t) \log \left(\frac{C(t)}{L(t)} \right) dt, \quad \rho > \lambda > 0, \quad (1)$$

where 0 is the arbitrary point in time when it makes decisions, ρ is the individual discount rate, $C(t)$ is aggregate consumption, and $L(t) = L_0 e^{\lambda t}$, $L_0 \equiv 1$, is population size (the mass of household members). Since each household member is endowed with one unit of time, $L(t)$ is the household's endowment of labor. The household's endowment of land is the constant Ω .

Let w and p denote, respectively, the price of labor and land services. In this setup the household supplies labor and land services inelastically and thus faces the flow budget constraint

$$\dot{A} = rA + wL + p\Omega - C, \quad (2)$$

where A is assets holding and r is the rate of return on assets. Denoting $c \equiv \dot{C}/C - \lambda$, the consumption plan that maximizes (1) subject to (2) consists of the Euler equation

$$r = \rho + c, \quad (3)$$

the budget constraint (2) and the usual boundary conditions.

3.2 Final producers

A competitive representative firm produces a final good Y that can be consumed, used to produce intermediate goods, invested in the improvement of the quality of existing intermediate goods, or invested in the creation of new intermediate goods. The final good is the numeraire so its price is

$P_Y \equiv 1$. To keep things simple, there is no physical capital.⁸ The production technology is

$$Y = \int_0^N X_i^\theta \left(Z_i^\alpha Z^{1-\alpha} \left(\frac{L}{N^{\eta_L}} \right)^\gamma \left(\frac{\Omega}{N^{\eta_\Omega}} \right)^{1-\gamma} \right)^{1-\theta} di, \quad 0 < \theta, \alpha, \gamma, \eta_L, \eta_\Omega < 1, \quad (4)$$

where N is the mass of non-durable intermediate goods, X_i is the quantity of good i , and L and Ω are, respectively, the services of labor and land purchased from the household.⁹ Quality is the good's ability to raise the productivity of L and Ω . More specifically, the contribution of good i to the productivity of the non-reproducible factors depends on its own quality, Z_i , and on average quality, $Z = \int_0^N (Z_j/N) dj$.¹⁰

The parameters η_L and η_Ω capture the degree of congestion (or rivalry¹¹) of the services of labor and land across intermediate goods. For $\eta_L = \eta_\Omega = 0$ there is no congestion, meaning that services of labor and land can be shared by all intermediate goods with no loss of productivity. This is a case of extreme economies of scope in the use of the services of the physical factor of production L and Ω that, as I show below, in the reduced-form representation of the production function in equilibrium manifest themselves as strong social increasing returns to product variety. At the opposite end of the spectrum, $\eta_L = \eta_\Omega = 1$ yields full congestion, where there are no economies of scope and therefore no social returns to variety. For $0 < \eta_L, \eta_\Omega < 1$ the technology features social returns to variety of degree less than one and social returns to quality of degree 1.

The first-order conditions for the profit maximization problem of the final producer yield that each intermediate producer faces the demand curve

$$X_i = \left(\frac{\theta}{P_i} \right)^{\frac{1}{1-\theta}} Z_i^\alpha Z^{1-\alpha} \left(\frac{L}{N^{\eta_L}} \right)^\gamma \left(\frac{\Omega}{N^{\eta_\Omega}} \right)^{1-\gamma}, \quad (5)$$

where P_i is the price of good i . Also, the final producer pays total compensation

$$\int_0^N P_i X_i di = \theta Y, \quad wL = \gamma(1-\theta)Y \quad \text{and} \quad p\Omega = (1-\gamma)(1-\theta)Y \quad (6)$$

to intermediate goods, labor and land suppliers, respectively.

⁸More precisely, there is no physical capital in the neoclassical sense of a homogenous, durable, intermediate good accumulated through foregone consumption. Instead, there are differentiated, non-durable, intermediate goods produced through foregone consumption. One can think of these goods as capital, albeit with 100% instantaneous depreciation. Introducing neoclassical physical capital complicates the analysis without adding insight.

⁹This representation implicitly imposes labor and land market clearing.

¹⁰This specification modifies the augmented Schumpeterian model developed by Aghion and Howitt (1998) to make it better suited to my purposes and yet leave the core mechanism essentially unchanged. The first modification is diminishing private returns to own quality, i.e., $\alpha < 1$. This allows me to work with symmetric equilibria that feature creative accumulation, whereby all incumbent firms do R&D, as opposed to creative destruction, whereby outsiders do R&D to replace the current incumbent. The second modification is that quality enters with exponent $1-\theta$, instead of 1, because my intermediate producers face a marginal cost of production in units of the final good, instead of a marginal cost in units of (physical) capital proportional to their quality level. Both approaches imply that quality enters the reduced-form version of (4) as augmenting the input in exogenous supply, which here is a Cobb-Douglas composite of labor and land.

¹¹Rivalry can be modeled by writing the labor and land inputs with a subscript i to capture that their services are assigned to the specific good i and cannot be shared with the other goods. The approach in the text is simpler.

3.3 Intermediate producers

The typical intermediate firm operates a technology that requires one unit of final output per unit of intermediate good and a fixed operating cost $\phi Z_i^\alpha Z^{1-\alpha}$, also in units of final output. The firm can increase quality according to the technology

$$\dot{Z}_i = I_i, \quad (7)$$

where I_i is R&D in units of final output. Using (5), the firm's gross profit (i.e., before R&D) is

$$\Pi_i = \left[(P_i - 1) \left(\frac{\theta}{P_i} \right)^{\frac{1}{1-\theta}} \left(\frac{L}{N^{\eta_L}} \right)^\gamma \left(\frac{\Omega}{N^{\eta_\Omega}} \right)^{1-\gamma} - \phi \right] Z_i^\alpha Z^{1-\alpha}. \quad (8)$$

The firm chooses the time path of price and R&D in order to maximize the price of its shares,

$$V_i(0) = \int_0^\infty e^{-\int_0^t r(s)ds} [\Pi_i(t) - I_i(t)] dt, \quad (9)$$

subject to (7) and (8), where r is the interest rate and 0 is the arbitrary point in time when the firm makes decisions. The firm takes average quality, Z , in (8) as given. The characterization of the firm's decision yields a symmetric equilibrium where

$$r = \alpha \frac{\Pi}{Z} \quad (10)$$

is the return to quality innovation (see the Appendix for the derivation) and α is now intuitively interpreted as the elasticity of the firm's gross profit with respect to its own quality.

New products are developed by entrepreneurs that set up new firms to serve the market. To start up activity an entrepreneur must sink $\beta Y/N$ units of final output (see Barro and Sala-i-Martin 2004, pp. 300-302, for arguments in support of this assumption). Because of this sunk cost, the new firm cannot supply an existing good in Bertrand competition with the incumbent monopolist but must instead introduce a new good that expands product variety. Entry is positive if the value of the new firm is equal to its setup cost, i.e., if $V_i = \beta Y/N$. Entrepreneurs finance entry by issuing equity and enter at the average quality level. The latter is a simplifying assumption that preserves symmetry of equilibrium at all times. The free-entry condition then yields the return to variety innovation (see the Appendix for the derivation)

$$r = \frac{N(\Pi - I)}{\beta Y} + \frac{\dot{Y}}{Y} - \frac{\dot{N}}{N}. \quad (11)$$

4 The economy's dynamics

This section focuses on the key allocation problem of the economy — the allocation of final output Y to consumption, production of intermediates and, when profitable, vertical and horizontal innovation — and derives the reduced-form representation of the resulting equilibrium dynamics. The representation yields an analytical solution for the economy's path.

4.1 General equilibrium

Intermediate producers set $P = 1/\theta$. Imposing symmetry in the production function (4) and using (5) to eliminate X yields the following reduced-form production function for final output:

$$Y = \theta^{\frac{2\theta}{1-\theta}} \cdot N^\sigma Z L^\gamma \Omega^{1-\gamma}, \quad \sigma \equiv 1 - \gamma\eta_L - (1 - \gamma)\eta_\Omega, \quad (12)$$

where the composite $N^\sigma Z$ represents the state of technology and σ is the degree of social returns to variety. Taking logs and time derivatives of (12) and subtracting population growth yields

$$y = \sigma n + z - (1 - \gamma)\lambda, \quad (13)$$

where $y \equiv \dot{Y}/Y - \lambda$, $n \equiv \dot{N}/N$ and $z \equiv \dot{Z}/Z$. In words, final output per capita growth is given by the growth rate of technology minus the *growth drag* due to the presence of land, which over time becomes relatively more and more scarce as population L grows exponentially at rate λ . Of course, if $\gamma = 1$ the drag disappears.

The key component of the model is the characterization of the incentives to vertical and horizontal innovation. As equations (10)-(11) and the definition of gross profit (8) show, the returns to both activities depend on the gross cash flow of the firm $X(P - 1)$ — i.e., revenues minus variable production costs — since this is the appropriate measure of profitability for firms that spread fixed costs, including the cost of developing innovations, over their own volume of sales. On closer inspection, moreover, one can see that both returns can be written as functions of the *quality-adjusted* gross cash flow of the firm. It is thus useful to define

$$x \equiv \frac{X(P - 1)}{Z} = \frac{\text{gross cash flow}}{\text{quality}}$$

and use it as the model's key measure of firm size.

Intermediate producers receive $N \cdot PX = \theta Y$ from the final producer. Consequently, $NX = \theta^2 Y$. Equation (12) then yields

$$x = \theta(1 - \theta) \frac{Y}{NZ} = (1 - \theta) \theta^{\frac{1+\theta}{1-\theta}} \cdot \frac{L^\gamma \Omega^{1-\gamma}}{N^{1-\sigma}}. \quad (14)$$

Substituting into (10) and (11) the returns to innovation become:

$$r = \alpha(x - \phi); \quad (15)$$

$$r = \frac{1}{\pi} \left(1 - \frac{\phi + z}{x} \right) + \frac{\dot{x}}{x} + z, \quad (16)$$

where to simplify notation I define

$$\pi \equiv \frac{\beta}{\theta(1 - \theta)} = \frac{\beta Y/N}{X(P - 1)} = \frac{\text{entry cost}}{\text{gross cash flow}}.$$

Expressions (15)-(16) capture the model's main property: firm-level decisions depend on the quality-adjusted cash flow, which is increasing in population and the land endowment because they

drive production of final goods and thereby demand for intermediate goods. It should be clear, thus, that from the viewpoint of the managers of incumbent firms and of the entrepreneurs that set up new firms the critical market size variable is expenditure on intermediates, θY . Recall, moreover, that consumption, production of intermediates and quality and variety innovation are all in units of the final good so that the resource allocation problem of this economy is the allocation across its alternative uses of the quantity Y produced according to (12). The following property characterizes the consumption flow that results from such allocation.

Proposition 1 *In equilibrium the economy's consumption ratio is*

$$\frac{C}{Y} = \begin{cases} (1 - \theta) \left[\theta \left(1 - \frac{\phi+z}{x} \right) + 1 \right] & n = 0 \quad z \geq 0 \\ (1 - \theta) [\theta (\rho - \lambda) \pi + 1] & n > 0 \quad z \geq 0 \end{cases}. \quad (17)$$

Proof. See the Appendix. ■

When entry is positive the fraction of final output that is consumed is constant throughout the transition as well as in steady state. When entry is zero, instead, the consumption ratio is increasing in firm size, x , and decreasing in the R&D intensity of the firm, $z = I/Z$, if positive. The reason is that incumbents earn rents that they distribute to the household as dividends. These rents increase with the size of the market (the numerator of x) because the fixed operating cost ϕZ (under symmetry) implies a falling unit production cost as the scale of operations of the firm rises. When entrants become active, these rents are competed away and the consumption ratio no longer rises with firm size.¹²

4.2 Horizontal and vertical innovation

Proposition 1 highlights that there exist corner solutions where one or both of the two R&D activities shut down. When entrants are active and the consumption ratio c is constant, the return to saving (3) reduces to $r = \rho + y$ and thus (16) yields

$$n = \begin{cases} \frac{1}{\pi} \left(1 - \frac{\phi+z}{x} \right) - \rho + \lambda & z > 0 \\ \frac{1}{\pi} \left(1 - \frac{\phi}{x} \right) - \rho + \lambda & z = 0 \end{cases}, \quad (18)$$

which says that there is a threshold of firm size below which entrants are not active ($n = 0$) because the return is too low. The value of the threshold depends on whether entrants anticipate that in the post-entry equilibrium $z > 0$ or $z = 0$ since it affects the net cash flow that they anticipate to earn. Similarly, the saving schedule (3), the reduced-form production function (12), the return to quality innovation (15) and the top line of (18) yield

$$z = \begin{cases} \alpha (x - \phi) - \rho + \lambda - \sigma n - \gamma \lambda & n > 0 \\ \alpha (x - \phi) - \rho + \lambda - (c - y) - \gamma \lambda & n = 0 \end{cases}, \quad (19)$$

¹²This property deserves a much more detailed discussion. To preserve this section's focus on solving analytically the model, I postpone it to the next section where I discuss the anatomy of the growth path of the economy.

which says that there is a threshold of firm size below which incumbents do not do quality R&D ($z = 0$) because the return is too low. The value of the threshold depends on whether $n > 0$ or $n = 0$ since it affects the return to innovation that they anticipate they must earn to deliver to their stockholders (the savers) their reservation rate of return on assets.

The interdependence of agents' activation decisions implies that the *sequence* in which the economy turns on the two innovation engines determines the shape of its transition path and the timing of the key events. It is useful to begin the analysis with a characterization of the equilibrium where both types of R&D are positive.

Proposition 2 *Let x_N denote the threshold of firm size for variety innovation and x_Z the threshold of firm size for quality innovation. Assume $\pi x > \sigma \ \forall x \geq \phi$, i.e., $\pi\phi > \sigma$. Then, for $x > \max\{x_Z, x_N\}$ the equilibrium rates of variety and quality innovation are:*

$$n = \frac{(1 - \alpha - (\rho - \lambda)\pi)x - (1 - \alpha)\phi + \rho - \lambda + \gamma\lambda}{\pi x - \sigma} > 0; \quad (20)$$

$$z = \frac{(x - \phi)\left(\alpha - \frac{\sigma}{\pi x}\right) - (1 - \sigma)(\rho - \lambda) - \gamma\lambda}{1 - \frac{\sigma}{\pi x}} > 0. \quad (21)$$

Proof. See the Appendix. ■

Thus, if firm size grows sufficiently large, the economy turns on both innovation engines. Setting aside for the moment the issue of stability, it is useful to characterize the steady state associated to such equilibrium.

Proposition 3 (The modern growth steady state) *Assume:*

$$\phi\pi > \frac{1}{\alpha} \left[1 - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right) \right] > 1.$$

Then in the region $x > \max\{x_Z, x_N\}$ there exists the steady state:

$$x^* = \frac{(1 - \alpha)\phi - \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma}\right)}{1 - \alpha - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma}\right)} > 0; \quad (22)$$

$$n^* = \frac{\gamma\lambda}{1 - \sigma} > 0; \quad (23)$$

$$z^* = \left[\frac{\alpha(\phi\pi - 1)}{1 - \alpha - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma}\right)} - 1 \right] \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right) > 0. \quad (24)$$

This steady state exhibits growth of final output per capita

$$y^* = \sigma n^* + z^* - (1 - \gamma)\lambda = \alpha(x^* - \phi) - \rho,$$

which is positive iff

$$\alpha \frac{(\phi\pi - 1) \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right)}{1 - \alpha - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right)} > \rho. \quad (25)$$

Proof. See the Appendix. ■

This proposition establishes conditions under which the economy possesses a steady state where both types of R&D are positive and the growth rate of final output per capita is constant and positive. Does the economy converge to such steady state?

4.3 Dynamics

The structure of incentives for innovation discussed above identifies conditions that yield two sequences of events.

Proposition 4 *There exists a combination of values of the parameters such that the thresholds x_N and x_Z are identical. There are thus two regimes, characterized by the order in which the economy activates the quality and the variety engines of growth.*

- **Dominant incentives for variety innovation.** For parameters such that

$$\gamma\lambda \left(\rho - \lambda + \frac{\theta + 1}{\theta(\rho - \lambda)\pi + 1} \right) > \alpha\phi \frac{(\rho - \lambda)\pi}{1 - (\rho - \lambda)\pi},$$

the ordering of the thresholds is $x_N < x_Z$, where

$$\begin{aligned} x_N &= \frac{\phi}{1 - (\rho - \lambda)\pi}; \\ x_Z &= \arg \text{solve} \left\{ (x - \phi) \left(\alpha - \frac{\sigma}{\pi x} \right) = (1 - \sigma)(\rho - \lambda) + \gamma\lambda \right\}. \end{aligned}$$

- **Dominant incentives for quality innovation.** For parameters such that

$$\gamma\lambda \left(\rho - \lambda + \frac{\theta + 1}{\theta(\rho - \lambda)\pi + 1} \right) < \alpha\phi \frac{(\rho - \lambda)\pi}{1 - (\rho - \lambda)\pi},$$

the ordering of the thresholds is $x_Z < x_N$, where

$$\begin{aligned} x_N &= \frac{(1 - \alpha)\phi - \rho + \lambda - \gamma\lambda}{1 - \alpha - (\rho - \lambda)\pi}; \\ x_Z &= \arg \text{solve} \left\{ \alpha(x - \phi) = \frac{1 + 1/\theta}{1 - \frac{\phi}{x} + 1/\theta} \gamma\lambda \right\}. \end{aligned}$$

Proof. See the Appendix. ■

The difference between the two cases is that in the variety-first case x_Z is the threshold for quality R&D given that the market *already* supports entry of new firms, whereas in the quality-first case it is the threshold for quality R&D given that the market *does not yet* support entry of new firms. Accordingly, in the variety-first case incumbents undertaking quality R&D compete for resources with entrants and face a constant reservation interest rate demanded by savers. In the quality-first case, instead, they do not compete for resources with entrants that are setting up new

firms but, because the free entry condition does not hold and they distribute to shareholders rents that grow with the size of the market, they face a reservation interest rate that reflects the growing consumption ratio. Similar reasoning applies to the threshold x_N . The following proposition provides the paper's main analytical result.

Proposition 5 *Let the economy's initial condition be*

$$x_0 = (1 - \theta) \theta^{\frac{1+\theta}{1-\theta}} \cdot \frac{L_0^\gamma \Omega^{1-\gamma}}{N_0^{1-\sigma}} < \min \{x_N, x_Z\}$$

and recall the steady-state firm size x^ characterized in equation (22). The two regimes then yield the following dynamics.*

- **The variety-first path to modern growth.** *The equilibrium law of motion of firm size is piecewise linear,*

$$\dot{x} = \begin{cases} \gamma\lambda x & x \leq x_N \\ \bar{\nu}(\bar{x}^* - x) & x_N < x \leq x_Z \\ \nu(x^* - x) & x > x_Z \end{cases}, \quad (26)$$

with coefficients,

$$\begin{aligned} \bar{\nu} &\equiv \frac{1-\sigma}{\pi} \left[1 - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1-\sigma} \right) \right], \\ \bar{x}^* &\equiv \frac{\phi}{1 - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1-\sigma} \right)}, \\ \nu &\equiv \frac{1-\sigma}{\pi} \left[1 - \alpha - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1-\sigma} \right) \right], \end{aligned}$$

and yields the explicit solution $x(t)$ illustrated in Figure 1: firm size follows an S-shaped path with inflection point at T_N and convergence from below to x^ , where*

$$T_N = \frac{1}{\gamma\lambda} \log \left(\frac{x_N}{x_0} \right) \quad (27)$$

is the date when the economy crosses the threshold x_N and turns on variety innovation and

$$T_Z = T_N + \frac{1}{\bar{\nu}} \log \left(\frac{\bar{x}^* - x_N}{\bar{x}^* - x_Z} \right). \quad (28)$$

is the date when the economy crosses the threshold x_Z and turns on quality innovation.

- **The quality-first path to modern growth.** *The equilibrium law of motion of firm size is piecewise linear,*

$$\dot{x} = \begin{cases} \gamma\lambda x & x \leq x_Z \\ \gamma\lambda x & x_Z < x \leq x_N \\ \nu(x^* - x) & x > x_N \end{cases}, \quad (29)$$

and yields the explicit solution $x(t)$ illustrated in Figure 2: firm size follows an S-shaped path

Figure 1: The time path of firm size in the variety-first case

Figure 2: The path of firm size in the quality-first case

with inflection point at T_N and convergence from below to x^* , where

$$T_Z = \frac{1}{\gamma\lambda} \log \left(\frac{x_Z}{x_0} \right) \quad (30)$$

is the date when the economy crosses the threshold x_Z and turns on quality innovation and

$$T_N = T_Z + \frac{1}{\gamma\lambda} \log \left(\frac{x_N}{x_Z} \right) = \frac{1}{\gamma\lambda} \log \left(\frac{x_N}{x_0} \right). \quad (31)$$

is the date when the economy crosses the threshold x_N and turns on variety innovation.

Proof. See the Appendix. ■

The difference between the two solutions for the last part of the equilibrium path is only in the time periods over which they hold, which are determined by the starting dates T_Z and T_N . The key difference between the two solutions, however, is that in the variety-first case premature market saturation that prevents the economy from reaching the phase of quality innovation is possible. This outcome is the red (lighter) path in Figure 1 that converges to the steady state $\bar{x}^*$.

4.4 Discussion

The reduced-form, state-space representation of this model consists of a pair of piece-wise linear differential equations in firm size x characterizing two possible scenarios. In one the incentives for horizontal innovation dominate and the economy crosses the threshold for variety innovation first. In the other the reverse is true: the incentives for vertical innovation dominate and the economy crosses the threshold for quality innovation first.

Proposition 4, specifically the expressions for the smaller threshold in each of the two cases, says that a finite threshold of firm size that activates one or the other innovation engine always exists. This means that as long as population growth is positive the economy *must* turn on Schumpeterian innovation. The intuition is that as long as the overall market for intermediate goods grows due to exogenous forces, firm size (i.e., profitability) grows and eventually must cross the threshold where one of the two engines of growth is turned on. This is, in essence, a no-arbitrage argument: as rents escalate the only force that can prevent agents from investing in activities aimed at capturing a share of them is infinite innovation costs or, equivalently, zero productivity of investment of final goods in variety and quality innovation (in the model's notation, $\beta \rightarrow \infty$ and $\alpha \rightarrow 0$).

Since innovation must start, the only question is *when* and what specific *sequence of events* unfolds. Proposition 4 says that the model's parameters space consists of two thick regions, one where variety innovation starts first, the other where quality innovation starts first. The two paths of firm size shown in Figures 1-2 generate drastically different economic histories.

5 Interpreting the model: three phases of growth

This section focuses on the model's predictions. It characterizes the economy's path in each case in terms of (a) within-phase behavior of key observables and (b) the timing of the transitions from

one phase to the next.

5.1 Anatomy of the transition: The variety-first case

Along the paths of the state variable x shown in Figures 1-2 the level of final production is given at any point in time by (12). That expression contains only the levels of the state variables N (product variety), Z (product quality) and L (population). Consequently, the path $Y(t)$ obtains from the paths $N(t)$, $Z(t)$ and $L(t)$. As argued, for simplicity the path of population is exogenous and exponential. Moreover, given initial conditions N_0 and Z_0 , and the solution $x(t)$, the paths of variety and quality are fully determined by equations (20) and (21). Although this procedure allows one to solve analytically for the paths $Y(t)$, $N(t)$, $Z(t)$ and $L(t)$, it is more insightful to characterize the evolution of the economy in terms of equations that express the relevant variables as functions of firm size x .

5.1.1 Final output, GDP and consumption

Proposition 1 shows that the allocation of final output across its alternative uses features a ratio C/Y that is increasing in firm size x when entrants are not active and constant when entrants are active. As argued, such behavior stems from static economies of scale that manifest themselves as efficiency gains in the production of intermediates.

To refine that intuition let G denote this economy's GDP. Subtracting the cost of intermediate production from the value of final production and using (14) yields

$$\frac{G}{Y} = (1 - \theta) \left[\theta \left(1 - \frac{\phi}{x} \right) + 1 \right]. \quad (32)$$

The term in brackets is increasing in x because the unit cost of production of the typical intermediate firm falls as its scale of operation rises. Taking logs and time derivatives of (32) yields

$$\frac{\dot{G}}{G} = \frac{\dot{Y}}{Y} + \xi(x) \frac{\dot{x}}{x}, \quad \xi(x) \equiv \frac{\theta\phi}{(1 + \theta)x - \theta\phi}, \quad (33)$$

where $\xi(x)$ is the elasticity of GDP with respect to firm size. This expression says that GDP growth is given by final output growth plus the contribution from efficiency gains in intermediate production due to firm size growth.

Equations (12)-(13) and (32)-(33) provide a complete characterization of output dynamics for this economy. Initially, final output grows only because of population growth, that is, $\dot{Y}/Y = \gamma\lambda$. Moreover, because consumption equals GDP, $\dot{C}/C = \dot{G}/G = [\xi(x) + 1]\gamma\lambda$. Thus, in the early Smithian phase with no Schumpeterian innovation, GDP and consumption growth are due solely to population growth and its amplification through static economies of scale. As the economy crosses the threshold x_N and enters the second phase, final output growth becomes $\dot{Y}/Y = \gamma\lambda + \sigma n(x)$, where the rate of variety growth $n(x)$ is given by the top line of (18). The third and final phase has both variety and quality innovation so that $\dot{Y}/Y = \gamma\lambda + \sigma n(x) + z(x)$, where $n(x)$ and $z(x)$ are given, respectively, by (20) and (21) in Proposition 2.

Figure 3: The growth rates of final output and GDP per capita as functions of firm size in the variety-first case

It is useful to summarize this characterization in terms of the growth rates of per-capita final output, GDP and consumption — y , g , and c , respectively — since these are the objects that the empirical literature typically discusses. Furthermore, it is useful to write these growth rates as the sum of a Schumpeterian innovation component that does not vanish in steady state and a Smithian component, due to static economies of scale, that vanishes in steady state:

$$y(x) = \begin{cases} -(1-\gamma)\lambda & \phi \leq x \leq x_N \\ \left(\frac{\gamma}{1-\sigma} - 1\right)\lambda - \frac{\sigma}{1-\sigma}\bar{\nu}\left(\frac{\bar{x}^*}{x} - 1\right) & x_N < x \leq x_Z \\ \alpha(x - \phi) - \rho & x > x_Z \end{cases} ;$$

$$g(x) = \begin{cases} ([\xi(x) + 1]\gamma - 1)\lambda & \phi \leq x \leq x_N \\ \left(\frac{\gamma}{1-\sigma} - 1\right)\lambda + \left[\xi(x) - \frac{\sigma}{1-\sigma}\right]\bar{\nu}\left(\frac{\bar{x}^*}{x} - 1\right) & x_N < x \leq x_Z \\ \alpha(x - \phi) - \rho + \xi(x)\nu\left(\frac{\bar{x}^*}{x} - 1\right) & x > x_Z \end{cases} ;$$

$$c(x) = \begin{cases} g(x) & \phi \leq x \leq x_N \\ y(x) & x > x_N \end{cases} .$$

Figures 3-4 illustrate these functions and the associated solutions for the growth rates.

The story that these equations tell is one where the economy starts out in a situation where there is no entry and firms earn rents. These rents grow with the size of the market and fuel GDP and consumption growth in excess of final output growth. Consequently, negative growth of final output per capita does not necessarily imply falling GDP and consumption per capita. In fact, it is possible to choose parameter values such that $[\xi(x) + 1]\gamma > 1$, meaning that GDP per capita grows

Figure 4: The growth rates of final output and GDP per capita as functions of time in the variety-first case

all the time.¹³ More generally, GDP per capita growth can start out negative and stay negative until the economy hits $x = x_N$ and turns on variety innovation. When that happens, the growth rate of GDP per capita starts rising and eventually turns positive if the contribution of product variety to final production is sufficiently strong. The growth rate of consumption per capita c , in contrast, drops to the growth rate of final output per capita, y , and remains below the growth rate of GDP per capita, g , until the end of the transition, where the constant firm size, $x = x^*$, yields a constant ratio between final output and GDP.

In the intermediate phase the tension between the exploitation of static economies of scale, that requires firm growth, and the exploitation of social returns to variety, that requires entry, results in a profile of GDP per capita growth that is always convex but can be increasing, U-shaped or decreasing in x throughout the interval $x_N < x \leq x_Z$. The parameter that drives these cases is the degree of social returns to variety σ . To avoid clutter, the figures illustrate only the second possibility, corresponding to situations where σ is sufficiently large that there exists a value of x where $dg(x)/dx$ becomes positive in the interval $x_N < x \leq x_Z$. If such value of x is larger than x_Z , which happens when the degree of social returns to variety σ is small, the function $g(x)$ is decreasing throughout the range $x_N < x \leq x_Z$ and the third case arises. This case is remarkable in that it says that the onset of systematic, profit-driven horizontal innovation is associated to a continuation of the slowdown of GDP per capita growth due to the gradual exhaustion of static economies of scale.

¹³Intuitively, this requires a restriction on the elasticity of output with respect to land in final production, i.e., $1 - \gamma \leq \theta(1 - (\rho - \lambda)\pi) / (1 + \theta)$.

5.1.2 Timing of events: the role of the fundamentals

The closed-form solution for the transition path provides analytical insight on the determinants of the timing of the key events in the economy's history. The expressions for T_N and T_Z in (27) and (28) reveal the following pattern.

- The activation of horizontal innovation occurs earlier, i.e., T_N is lower, in economies where the ratio x_N/x_0 is lower.
- The activation of vertical innovation occurs if and only if the steady-state firm size $\bar{x}^*$ associated to the variety-driven phase is smaller than the threshold for quality R&D x_Z .
- Given T_N , and conditional on $\bar{x}^* > x_Z$, the activation of vertical innovation occurs earlier, i.e., T_Z is lower, in economies where:
 - the steady-state firm size $\bar{x}^*$ is larger;
 - convergence in the variety-driven phase is faster, i.e., where $\bar{\nu}$ is higher;
 - the threshold for quality R&D x_Z is smaller.

Checking the expressions for x_0 , $\bar{\nu}$, x_N , x_Z , $\bar{x}^*$ provides further detail.

Proposition 6 *The date of activation of variety innovation, T_N , is:*

- *decreasing in the initial population L_0 , the land endowment Ω , the population growth rate λ and the elasticity of output with respect to labor γ ;*
- *increasing in the fixed operating cost ϕ , the sunk entry cost β and the discount rate ρ ;*
- *independent of the elasticity of gross profit with respect to own quality α and the degree of social returns to variety σ .*

The duration of the phase with variety innovation only, $T_Z - T_N$, is:

- *decreasing in the fixed operating cost ϕ and the elasticity of gross profit with respect to own quality α ;*
- *increasing in the degree of social returns to variety σ ;*
- *depends ambiguously on the population growth rate λ , the elasticity of output with respect to labor γ , the sunk entry cost β and the discount rate ρ ;*
- *independent of the initial population L_0 , the land endowment Ω .*

This characterization identifies factors that explain why some economies take off earlier than others — defining the take-off as the onset of systematic, profit-driven innovation — and factors that explain why some economies experience a faster transition than others to the ultimate phase with both innovation engines turned on and convergence to sustained growth. Moreover, it provides a novel insight on why some economies might fail to reach the modern growth phase: they might fail to turn on vertical innovation due to premature market saturation.

5.1.3 A closer look at consumption-saving behavior and factor remuneration

It is useful to examine the behavior of the ratio of consumption to GDP, which can be written

$$\frac{C}{G} = \frac{C}{Y} \frac{Y}{G} = \begin{cases} \frac{\theta(1-\frac{\phi+z}{x})+1}{\theta(1-\frac{\phi}{x})+1} & n = 0 \quad z \geq 0 \\ \frac{\theta(\rho-\lambda)\pi+1}{\theta(1-\frac{\phi}{x})+1} & n > 0 \quad z \geq 0 \end{cases}.$$

When entrants are not active ($n = 0$), the ratio is 1 if there is no vertical innovation ($z = 0$) because in that case the economy makes no investment and thus needs no saving. If instead there is vertical innovation ($z > 0$), the ratio is less than 1 and decreasing in z , because faster quality growth requires more investment, and increasing in x , because firm size growth leads to falling unit costs in intermediate production. When entrants are active ($n > 0$), the ratio is independent of z and decreasing in x . One can summarize these observations as follows:

$$\frac{C}{Y} = \begin{cases} (1-\theta) \left[\theta \left(1 - \frac{\phi}{x} \right) + 1 \right] & \phi \leq x \leq x_N \\ (1-\theta) [\theta(\rho-\lambda)\pi+1] & x > x_N \end{cases};$$

$$\frac{C}{G} = \begin{cases} 1 & \phi \leq x \leq x_N \\ \frac{\theta(\rho-\lambda)\pi+1}{\theta(1-\frac{\phi}{x})+1} & x > x_N \end{cases}.$$

The first expression, reproduced from Proposition 1 for convenience, captures the property already discussed that static efficiency gains drive consumption growth in excess of final output growth in the Smithian phase of the transition. The second expression confirms that such consumption growth comes from efficiency gains in intermediate production that raise GDP. The fact that the ratio of consumption to GDP is decreasing in firm size x when entrants are active captures the property that after the onset of systematic, profit-driven innovation the economy's investment share rises throughout the transition to the steady state x^* .

Associated to this pattern of consumption-saving, there is a pattern of factor remuneration driven by the following property: factors that earn a flow of payments proportional to final output Y earn a share of GDP that is decreasing in firm size x and is thus decreasing over time throughout the transition. As shown in Section 2, the three factors that enter the production technology (4) — labor, land and intermediate goods — belong to this category. So, if these factors earn falling shares of GDP over time, what factor earns a rising share of GDP? The answer is that throughout

the transformation of this economy what rises is the share of GDP earned by firms in the form of gross profits. Specifically, equation (8) and (32) yield

$$\frac{\text{gross profits}}{\text{GDP}} = \frac{N\Pi}{G} = \frac{\theta Y - N(X + \phi Z)}{Y - N(X + \phi Z)} = \frac{\theta \left(1 - \frac{\phi}{x}\right)}{\theta \left(1 - \frac{\phi}{x}\right) + 1}.$$

In the initial phase with no innovation, these profits are distributed to shareholders and consumed. Only after crossing the threshold for profitable innovation the economy exhibits saving and investment, resulting in a falling consumption share. More importantly, once it kicks in, the free entry condition implies that the rising profits are not escalating pure rents and, more important, that they are reinvested in innovation, thus driving the economy's growth acceleration.

5.2 Anatomy of the transition: The quality-first case

The quality-first case differs from the variety-first case in the intermediate phase and in the timing of its beginning and end. After the economy crosses the threshold x_Z , the growth rate of final output is $\dot{Y}/Y = \gamma\lambda + z$, where z is given by the bottom line of (19). Because that expression contains the growth rate of the consumption output ratio, I cannot just substitute terms to express the growth rates of Y , G and C as functions of x . However, I can use equation (17) to construct policy functions that provide the information needed to characterize the equilibrium path. The details are in the Appendix. Here it suffices to note that in the interval $x_Z < x < x_N$ there exists a function $z(x)$ that is increasing, convex and starts out with zero derivative at x_Z . Summarizing, the growth rates of final output, GDP and consumption per capita are:

$$y(x) = \begin{cases} -(1-\gamma)\lambda & 0 \leq x < x_Z \\ z(x) - (1-\gamma)\lambda & x_Z < x < x_N \\ \alpha(x-\phi) - \rho & x > x_N \end{cases} ;$$

$$g(x) = \begin{cases} ([\xi(x)+1]\gamma-1)\lambda & \phi \leq x \leq x_Z \\ ([\xi(x)+1]\gamma-1)\lambda + z(x) & x_Z < x \leq x_N \\ \alpha(x-\phi) - \rho + \xi(x)\nu\left(\frac{x^*}{x} - 1\right) & x > x_N \end{cases} ;$$

$$c(x) = \begin{cases} g(x) & \phi \leq x \leq x_N \\ y(x) & x > x_N \end{cases} .$$

Figures 5-6 illustrate these functions and the associated solutions for the growth rates.

Along this path the rate of innovation exhibits explosive behavior because firms start undertaking quality R&D when they are still earning escalating rents driven by aggregate market growth due to population growth. As in the previous case, these escalating rents fuel consumption growth in excess of final output growth. Moreover, since consumption per capita growth hits its minimum at $x = x_Z$, it is possible to choose parameter values such that GDP per capita growth is positive for all $\phi \leq x \leq x_Z$ and, consequently, that it grows all the time despite the fact that final output

Figure 5: The growth rates of final output and GDP per capita as functions of firm size in the quality-first case

Figure 6: The growth rates of final output and GDP per capita as functions of time in the quality-first case

per capita growth is initially negative. More generally, the economy can experience a period of negative GDP per capita growth until it hits $x = x_Z$. When that happens, the growth rate starts to rise and eventually turns positive. An interesting feature of this case is that, because the function $z(x)$ starts out with zero derivative at x_Z , it must be the case that the time profile of the growth rate of GDP is convex and decreasing at the onset of quality innovation. The reason is that the initial contribution of quality growth cannot overcome the gradual exhaustion of Smithian static economies of scale since it follows a very shallow time path.

The expressions for T_Z and T_N in (30) and (31) yield the following pattern for the timing of events and the role of the fundamentals.

- The activation of vertical innovation occurs earlier, i.e., T_Z is lower, in economies where the ratio x_Z/x_0 is lower.
- Given T_Z , the activation of horizontal innovation occurs earlier, i.e., T_N is lower, in economies where the threshold for variety R&D x_N is smaller.

Checking the expressions for x_0 , x_Z , x_N yields further detail.

Proposition 7 *The date of activation of quality innovation, T_Z , is:*

- *decreasing in the initial population L_0 , the land endowment Ω , the fixed operating cost ϕ and the elasticity of gross profit with respect to own quality α ;*
- *depends ambiguously on the population growth rate λ and the elasticity of output with respect to labor γ ;*
- *increasing in the discount rate ρ ;*
- *independent of the sunk entry cost β and the degree of social returns to variety σ .*

The duration of the phase with quality innovation only, $T_N - T_Z$, is:

- *increasing in the elasticity of gross profit with respect to own quality α and the sunk entry cost β ;*
- *depends ambiguously on the population growth rate λ , the elasticity of output with respect to labor γ , the fixed operating cost ϕ and the discount rate ρ ;*
- *independent of the initial population L_0 , the land endowment Ω and the degree of social returns to variety σ .*

As in the previous case, this characterization identifies factors that explain why some economies take off earlier than others and factors that explain why some economies experience a faster transition than others to the ultimate phase with sustained, modern growth.

5.3 Bringing it all together: When does the take-off occur?

The initial history of this economy is one of growth of GDP and consumption per capita driven by the amplification of population growth — more generally, aggregate market size growth driven by exogenous forces — through the exploitation of static economies of scale. This process of Smithian growth has been highlighted by many writers (e.g., Jones 1988, Mokyr 2005, 2010). The multiplier of population growth in the expressions above, the term $[\xi(x) + 1]\gamma - 1$, has a theoretical range of $(\gamma(\theta + 1) - 1, \gamma - 1)$ for $x \geq \phi$ so that even if one were to choose parameters that make it positive in the interval $\phi \leq x \leq \min\{x_N, x_Z\}$, it would yield a growth rate of GDP per capita equal to a fraction of the rate of population growth. Given that historically population growth rates prior to the Industrial Revolution were of the order of 0.1%, the model predicts very low growth rates of GDP and consumption per capita for the period.

How long does this stage of low growth last? Recall that the central message of Proposition 4 is that because population growth is positive at all times the economy *must* turn on Schumpeterian innovation and the only issue is which type it turns on first. Recall also that equations (27) and (30) differ only by the value of the threshold that the economy hits first. Consequently, it is convenient to define a generic value

$$x_T \equiv \min \left\{ \frac{\phi}{1 - (\rho - \lambda)\pi}, \arg \text{solve} \left\{ \alpha(x - \phi) = \frac{1 + 1/\theta}{1 - \frac{\phi}{x} + 1/\theta} \gamma \lambda \right\} \right\}$$

and think of the take-off date as $T = \frac{1}{\gamma\lambda} \ln(x_T/x_0)$. These expressions identify two main forces driving the duration of the initial phase. The first is the contribution of population growth to the growth of final output, $\gamma\lambda$. The second is the gap between the initial condition and the threshold where the economy activates innovation, x_T/x_0 . Using the definition of x_0 , the expression for the take-off time becomes

$$T = \frac{1}{\gamma\lambda} \ln \left(\frac{x_T}{x_0} \right) = \frac{1}{\gamma\lambda} \ln \left(\frac{x_T}{\theta(1 - \theta) \theta^{\frac{2\theta}{1-\theta}} \cdot \frac{L_0^\gamma \Omega^{1-\gamma}}{N_0^{1-\sigma}}} \right)$$

and identifies two additional sets of determinants. Technological and preference parameters drive the cost-benefit calculation underlying the activation decision, that is, the term x_T . The land endowment and the initial values of population and of the mass of intermediate firms/products do not enter this calculation; they show up only in the denominator as the determinants of the initial state of the economy.

Suppose that the economy has an initial value $x_0 = x_T/2$, that is, in order to cross the threshold that activates innovation, firm size has to double. Suppose also that $\gamma = 0.8$ and $\lambda = 0.001 = 0.1\%$. The expression above then shows that starting at time 0, the take-off time is $T = \ln 2 / (0.08\%) = (69.3\%) / (0.08\%) = 866.25$, or, using the “rule of 70” that approximates $\ln 2 = 70\%$, $T = \ln 2 / (0.08\%) = (70\%) / (0.08\%) = 875$.¹⁴ Thus, an economy with output elasticity

¹⁴The well-known and often used “rule of 72” that approximates $\ln 2 \simeq 72\%$ would yield $T = \ln 2 / (0.08\%) = (72\%) / (0.08\%) = 900$.

with respect to labor of 0.8 and whose population grows at 0.1% per year takes approximately 875 years to double its firm size. Note that such an economy experiences an increase in population given by $L(T) = 2^{1/\gamma} \cdot L_0 = 2.38 \cdot L_0$.

I have set up this example assuming $x_T/x_0 = 2$ because it simplifies the calculation by exploiting well-known heuristics. The expression for T , however, says much more about this ratio. Recall that the threshold x_T depends on technological and preference parameters and that L_0 , Ω and N_0 enter only in the determination of x_0 at the denominator. A key determinant of the take-off time is thus the initial fragmentation of the aggregate market for intermediate goods in submarkets and whether such fragmentation comes with little or large gains in productivity via social returns to variety. The definition $\sigma \equiv 1 - \gamma\eta_L - (1 - \gamma)\eta_\Omega$ relates such social returns to variety to the deeper congestion parameters that characterize the model.

There are thus several channels through which institutions and other social factors can enter the determination of the take-off time T . An economy with a larger population takes off faster only in the trivial *ceteris paribus* sense that the comparative statics effect of L_0 on x_0 is positive. What really matters in the theory, however, is how such an economy differs in terms of firm size from one with a smaller population. Once this is taken into account, what the model says is that the take-off time depends on a collection of determinants, including the availability of other factors of production (here land, more generally, resources, including exhaustible and/or renewable) and on how the underlying production structure determines congestion in the use of all factors of production across intermediate goods.

6 Other prominent features of the theory

The model has relatively few, standard ingredients and yet produces a rich set of results. Following are some properties worth emphasizing in a separate discussion to bring in even sharper relief what this paper's approach contributes to the literature.

Remark 8 *Prior to the onset of profit-driven systematic innovation, static economies of scale in intermediate production deliver income per capita growth in periods of population expansion. Such Smithian growth, however, is not self-sustaining and eventually must vanish.*

The best way to see this is to set parameters such that both $x_Z \rightarrow \infty$ and $x_N \rightarrow \infty$ (i.e., $\beta \rightarrow \infty$ and $\alpha \rightarrow 0$) so that innovation never takes hold. It then follows that asymptotically $[\xi(x) + 1]\gamma\lambda \rightarrow \gamma\lambda$ so that GDP per capita and consumption per capita growth converge to $-(1 - \gamma)\lambda$. This property is important because the historical record provides abundant evidence of sporadic bursts of income per capita growth, often associated to bursts of population growth. The main characteristics of these episodes is that they all eventually run out of steam and fizzled out. The model's key mechanism fits such pattern: population growth per se cannot give rise to self-sustaining growth of income per capita. As historians have stressed on multiple occasions (e.g., Jones 1988, Easterlin

1996, Mokyr 2005, 2010), the key to the growth acceleration that the world experienced in the 18th and 19th centuries is that it was associated to a qualitative transition to a different mode of growth, one based on sustained profit-driven innovation.¹⁵

Remark 9 *Changes in fundamentals that result in an earlier take-off date do not necessarily result in immediate take-off.*

This property sounds obvious but, on reflection, highlights something that the current debate on the timing of the Industrial Revolution seems to ignore: institutional changes that favor innovation do not result in immediate take-off if the economy has not yet matured the other necessary conditions for doing so. Specifically, an economy that at time t experiences an improvement in the business environment that results in lower thresholds x_Z and x_N does not take off at time t if $x(t) < \min\{x_Z, x_N\}$. In other words, an economy that at the time of the favorable institutional change has not yet achieved the required firm size *has to wait a shorter time to take off but does not take off immediately*. The current debate seems to take for granted that the response should be immediate (see, e.g., Mokyr 2005, 2010, Mokyr and Voth 2010, Galor 2005, 2011), probably because most of the models that deal with the issue postulate economies that need to be shocked out of a steady state with no growth.

Remark 10 *When the economy turns on quality innovation first, it exhibits explosive growth that ends in finite time.*

When firms start investing in quality innovation but the free entry condition does not yet apply, the dynamics replicate the special case of endogenous growth models driven by vertical innovation with a fixed number of products and exponential population growth. That is, it replicates the *explosive* growth that has been long considered problematic in first-generation models. This model, however, does not impose arbitrarily that product variety expansion is never operational so that at most a *finite* period of faster than exponential quality growth is possible. Explosive growth due to the scale effect, in other words, is *not* an inherent property of the theory. Rather, in first-generation models it is an artifact of the implausible assumption of fixed product variety — i.e., *infinite* entry costs — that prevents entry from competing away escalating rents.

Remark 11 *If the economy turns on variety innovation first, it can fail to cross the threshold for quality innovation.*

¹⁵Jones (1988), in particular, talks about a shift from Smithian to Promethean growth to emphasize the role of knowledge accumulation in the modern era. I follow Mokyr (2005, 2010) and talk about a shift from Smithian to Schumpeterian growth to emphasize the link between the historians' perspective and the economists' modern theory of innovation.

This property reinforces the previous observation about the importance of entry in competing away incumbents' rents. Not only entry tames explosive quality growth, it can also prevent the economy from reaching the stage where incumbents find profitable to improve their own products. Specifically, if variety innovation starts first and $\bar{x}^* \leq x_Z$, then $T_Z \rightarrow \infty$, which means that the dissipation of rents due to product proliferation is so strong that the economy stabilizes the value of firm size before it crosses the threshold for quality innovation.

Remark 12 *The steady-state mass of firms is not proportional to population but, rather, is a generic power function of population.*

Recall that in steady state firm size is constant. Accordingly, (14) yields

$$N^* = \left[(1 - \theta) \theta^{\frac{1+\theta}{1-\theta}} \cdot \frac{L^\gamma \Omega^{1-\gamma}}{x^*} \right]^{\frac{1}{1-\sigma}},$$

where x^* is independent of L and Ω ; see (22). Eliminating the scale effect through product proliferation, therefore, does *not* require the knife-edge assumption

$$N = (\text{constant}) \cdot L^\varkappa, \quad \varkappa = 1,$$

as often claimed (see, e.g., Jones 2005). Rather, the theory says

$$\varkappa = \frac{\gamma}{1 - \sigma} \lesseqgtr 1.$$

To get $\varkappa = 1$ one needs to assume either (a) $\gamma = 1$ (no land) and $\sigma = 0$ (no love of variety in production) or (b) $\gamma = 1 - \sigma \Rightarrow 1 - \gamma = \sigma$. Case (a) consists of simplifying assumptions that some of the early models imposed for convenience but that are not necessary features of the theory. Case (b) sets social returns to variety equal to the elasticity of output with respect to land. Recalling that $\sigma = 1 - \gamma\eta_L - (1 - \gamma)\eta_\Omega$ one sees that this is a special case requiring either that there is full congestion of labor ($\eta_L = 1$) associated to no congestion of land ($\eta_\Omega = 0$), or that somehow congestion of labor and land work out to $\gamma(1 - \eta_L) = (1 - \gamma)\eta_\Omega$.

Remark 13 *If the economy enters the ultimate phase with both variety and quality innovation, population expansion is no longer needed to pull income per capita growth. Indeed, the population growth rate can fall to zero with income per capita growth remaining positive.*

As the economy converges to the steady state, GDP per capita and consumption per capita growth converge to $g^* = \alpha(x^* - \phi) - \rho$, which is positive for $\alpha(x^* - \phi) > \rho$. The key to this expression is that meeting the condition for positive GDP per capita growth does not require special assumptions on population growth. Indeed, one can see from expression (22) that population growth λ can be zero (or even negative) without compromising the model's ability to deliver endogenous steady-state growth. This property is more important than it appears: it says that a burst of

population growth provides a *window of opportunity* that the economy can exploit to transition from its initial state with no innovation to the final state with endogenous, innovation-driven growth that does not require continuous market expansion due to exogenous forces.

The expression reveals something else as well: the effect of population growth on GDP per capita growth depends on the same condition that drives the steady-state relation between the mass of firms N and population size L . Specifically, x^* is increasing in $\rho - \lambda + \frac{\gamma\lambda}{1-\sigma}$ and therefore increasing in λ for $\gamma > 1 - \sigma$, independent of λ for $\gamma = 1 - \sigma$ and decreasing in λ for $\gamma < 1 - \sigma$. Recalling that $\sigma = 1 - \gamma\eta_L - (1 - \gamma)\eta_\Omega$ one sees that what drives the model's predictions about the effect of *exogenous* population growth on economic growth are the assumptions on congestion/rivalry of the services of the factors of production L and Ω across intermediate goods.¹⁶

7 Conclusion

This paper has proposed a theory of the emergence of modern Schumpeterian growth as the result of firms' and entrepreneurs' response to Smithian market expansion. The theory makes detailed predictions about the transition to innovation-driven growth, especially about the qualitative differences due to the timing and sequence of events.

As stated in the Introduction, to keep things simple, the paper takes no stand on early history or the demographic transition. The assumption of constant population growth is a simplification that, while convenient in deriving analytical results, deserves further scrutiny. I do not pursue this point here for reasons of space. I do so in related work in progress (Peretto 2013), where I argue that the Schumpeterian perspective adds an important dimension to our understanding of the Industrial Revolution. Moreover, population growth can be seen as just a stand-in force for exogenous market expansion. Alternatives to exogenous population growth that I have explored are: exogenous disembodied technological change; exogenous growth of the resource endowment (e.g., discovery

¹⁶In fact, things are even more interesting than this because the effects of population growth depend also on the assumptions one makes on preferences. In this paper I use the usual Benthamite specification that adds up utility across family members. Alternative assumptions are feasible. For example, one could modify (1) as follows

$$U(0) = \int_0^\infty e^{-\rho t} L^\mu(t) \log\left(\frac{C(t)}{L(t)}\right) dt, \quad 0 \leq \mu \leq 1$$

resulting into an effective discount rate of $\rho - \mu\lambda$ that captures the range of attitudes going from the case of no preference for family members ($\mu = 0$) to the Benthamite case discussed in the text ($\mu = 1$). With this specification, the expression for steady-state firm size becomes

$$x^* = \frac{(1 - \alpha)\phi - \left(\rho - \mu\lambda + \frac{\gamma\lambda}{1-\sigma}\right)}{1 - \alpha - \pi\left(\rho - \mu\lambda + \frac{\gamma\lambda}{1-\sigma}\right)}$$

while

$$g^* = \alpha(x^* - \phi) - \rho + \mu\lambda - \lambda,$$

which emphasizes how different assumptions on preferences result in different conclusions on the effect of population growth on economic growth. In particular, there are a direct negative effect due to $\mu < 1$ and an indirect effect through firm size that depends on $\gamma/\mu \lesseqgtr 1 - \sigma$.

and opening up of new land); growth of embodied knowledge through “natural” curiosity and/or learning by being/doing effects (this is essentially what drives UGT) — an interesting specification of this mechanism would be human capital accumulation through studying/research by the elites and through experience and/or experimentation on the job by common folks.

Such specifications of the forces that drive market size growth yield qualitatively similar results. They differ from exogenous population growth in that, absent a Malthusian mechanism, they all yield *rising* income per capita throughout the transition. The model thus formulated, therefore, would mask one of the key aspects of the Schumpeterian mechanism that I studied in this paper. Namely, that there is a fundamental difference between the forces driving income per capita and the forces driving profit per firm. It is the latter that drive the phase transition from the regime with no profit-driven innovation to that with profit-driven innovation. More importantly, the activation of the Schumpeterian engine of endogenous growth may well require that agents tolerate temporarily falling income and/or consumption per capita as population growth builds up the economy to the point where the scale of operations of firms is sufficiently large.

There are other, less prominent, differences between this paper and UGT that are worth highlighting. UGT considers only land as the factor that induces diminishing returns to labor and, typically, has no market for land. The second feature implies that UGT lacks a scarcity price signal and consequently has limited applicability to broader issues arising from the interactions of technology, demography and natural resources. Likewise, the first feature leaves out sources of scarcity that play an important role in the debate on the future growth potential of modern economies. This paper, in contrast, develops a framework that has a scarcity price signal and that extends seamlessly to the case of exhaustible and renewable natural resources. The ambition is to apply the ideas developed here to a broader class of questions. For an example, see Peretto (2013).

Another difference is that UGT typically has no consumption/saving decision. This model does, and thus applies a framework that assigns an important role to the financial market in channelling resources from savers to the agents with the investment projects in need of funding. In this perspective, this paper embeds the question of the long-run acceleration of the economy at the time of the Industrial Revolution in a more “traditional” macroeconomic framework.

None of these observations are criticisms of UGT. On the contrary, they are meant to highlight the *complementarity* between UGT and the Schumpeterian approach proposed in this paper. UGT has accomplished much in advancing our understanding of issues that for a long time have resisted our best analytical efforts. Much remains to be done, however, and the Schumpeterian approach to endogenous innovation allows us to make further progress in the areas that UGT had to simplify to keep the models tractable.

8 Appendix

8.1 Derivation of the returns to quality and variety innovation

The usual method of obtaining first-order conditions is to write the Hamiltonian for the optimal control problem of the firm. This derivation highlights the intuition. The firm undertakes R&D up to the point where the shadow value of the quality innovation, q_i , is equal to its cost,

$$1 = q_i \Leftrightarrow I_i > 0. \quad (34)$$

Since the innovation is implemented in-house, its benefits are determined by the marginal profit it generates. Thus, the return to the innovation must satisfy the arbitrage condition

$$r = \frac{\partial \Pi_i}{\partial Z_i} \frac{1}{q_i} + \frac{\dot{q}_i}{q_i}. \quad (35)$$

To calculate the marginal profit, observe that the firm's problem is separable in the price and investment decisions. Facing the isoelastic demand (5) and a marginal cost of production equal to one, the firm sets $P_i = 1/\theta$. Substituting this result into (8), differentiating with respect to Z_i , substituting into (35) and imposing symmetry yields (10).

To obtain the return to variety innovation, observe first that the value of the new firm is given by (9) because the post-entry profit flow that accrues to the entrant in is given by (8). Entrepreneurs undertake R&D up to the point where the value of the variety innovation, V_i , is equal to its cost,

$$\beta Y/N = V_i \Leftrightarrow \dot{N} > 0. \quad (36)$$

Taking logs and time derivatives of the value of the firm yields the rate of return to entry as

$$r = \frac{\Pi_i - I_i}{V_i} + \frac{\dot{V}_i}{V_i}. \quad (37)$$

Using the free-entry condition (36) and imposing symmetry yields (11).

8.2 Proof of Proposition 1

When $n > 0$ assets market equilibrium requires

$$A = NV = \beta Y, \quad (38)$$

which says that the wealth ratio A/Y is constant. This result and the saving schedule (3) allow me to rewrite the household budget (2) as the following unstable differential equation in C/Y ,

$$0 = \rho - \lambda + \left(\frac{\dot{C}}{C} - \frac{\dot{Y}}{Y} \right) + \frac{1 - \theta - (C/Y)}{\beta},$$

which says that to satisfy the transversality condition C/Y jumps to the constant value $(\rho - \lambda)\beta + 1 - \theta$. Using the definition of π yields the bottom line of (17).

When $n = 0$ assets market equilibrium still requires $A = NV$ but it is no longer true that $V = \beta Y/N$ since by definition the free-entry condition does not hold. This means that the wealth ratio A/Y is not constant. However, (37) holds since it is the arbitrage condition on equity holding that characterizes the value of an existing firm regardless of how it came into existence in the first place. Imposing symmetry and inserting (8), (37) and (38) into the household budget (2) yields

$$0 = N [(1/\theta - 1) X - \phi Z - I] + (1 - \theta) Y - C.$$

The definition of x , the R&D technology (7), and the fact that $NX = \theta^2 Y$, allow me to rewrite this expression as the top line of (17).

8.3 Proof of Proposition 2

The key to the proof is that it looks for a Nash equilibrium where both R&D activities yield a rate of a return that matches the reservation rate of return on saving of the household. Consider (18) and (19). In (z, n) space these are two negatively sloped straight lines – the quality innovation line and the variety innovation line, respectively – with intersections with the axes that depend positively on x . Solving the top lines of (18) and (19) for n and z yields (20) and (21). This solution is a stable Nash equilibrium if the variety innovation line (18) is flatter than the quality innovation line (19), that is, if $\pi x > \sigma$, which is surely true under $\pi\phi > \sigma$ since this model requires $x \geq \phi$.

8.4 Proof of Proposition 3

The definition of firm size (14) and the reduced-form production function (12) yield $\dot{x}/x = \gamma\lambda - (1 - \sigma)n$. Setting $\dot{x} = 0$ yields (23). Inserting (23) in (20) yields (22). Inserting (22) in (21) yields (24). The values x^* and z^* are positive if:

$$\begin{aligned} 1 - \alpha &> \frac{1}{\phi} \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right); \\ 1 - \alpha &> \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right); \\ \frac{\alpha(\phi\pi - 1)}{1 - \alpha - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right)} &> 1. \end{aligned}$$

Observing that the third inequality can hold only if $\phi\pi > 1$, which implies

$$1 - \frac{1}{\phi} \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right) > 1 - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1 - \sigma} \right),$$

yields the two conditions in the text of the proposition. Finally, noticing that C/Y is constant, one can use the Euler equation (3) to write

$$\begin{aligned} y^* &= r^* - \rho \\ &= \alpha(x^* - \phi) - \rho > 0 \end{aligned}$$

iff (25) holds.

8.5 Proof of Proposition 4

The proof is a generalization of that of Proposition 2: it looks for a Nash equilibrium where either at least one of the two R&D activities yields a rate of a return that matches the reservation rate of return on saving of the household or there is no TFP growth. As seen in Proposition 2, the Nash equilibrium features both vertical and horizontal R&D if they yield equal rates of return. Now consider (18) and (19) in (z, n) space and suppose that initially x is so small that both lines lie entirely below the origin. This configuration arises when both vertical and horizontal R&D yield a rate of return lower than what the household demands to postpone consumption and therefore the economy is in an equilibrium with zero TFP growth. As x grows two sequences of events are possible.

- The variety innovation line (18) reaches the origin and enters the positive quadrant before the quality innovation line (19). In this case, the economy crosses the firm size threshold that activates horizontal innovation while agents anticipate zero vertical innovation. That is, agents anticipate $z = 0$ and therefore (18) says that $n > 0$ if

$$x > x_N \equiv \frac{\phi}{1 - (\rho - \lambda)\pi} > \phi.$$

As x keeps growing, the quality innovation line (19) enters the positive quadrant and, if it catches up and overtakes the variety innovation line (18), the Nash equilibrium with both vertical and horizontal R&D takes hold. Specifically, given that along this path agents anticipate $n > 0$ at the switch point, (21) says that $z > 0$ if

$$(x - \phi) \left(\alpha - \frac{\sigma}{\pi x} \right) > (1 - \sigma)(\rho - \lambda) + \gamma\lambda.$$

The left-hand side starts out at zero for $x = \phi$ and is monotonically increasing in x . The inequality thus identifies a unique value

$$x_Z = \arg \text{solve} \left\{ (x - \phi) \left(\alpha - \frac{\sigma}{\pi x} \right) = (1 - \sigma)(\rho - \lambda) + \gamma\lambda \right\} > x_N.$$

- The quality innovation line (18) reaches the origin and enters the positive quadrant before the variety innovation line (20). In this case, the economy crosses the firm size threshold that activates vertical innovation while agents anticipate zero variety innovation, that is, agents anticipate $n = 0$ and $z > 0$ if $\alpha(x - \phi) > \rho - \lambda + \gamma\lambda + (c - y)$, where the value of $(c - y)$ comes from log-differentiating (17) with respect to time. It is useful to think of $z = 0$ if $\alpha(x - \phi) \leq \rho - \lambda + \gamma\lambda + (c - y)$ so that I can compute $(c - y)$ from (17) under $z = 0$. Recalling that I am working in the region where $\dot{x}/x = \gamma\lambda$, I then have that $z = 0$ for

$$\alpha(x - \phi) \leq \rho - \lambda + \frac{1 + 1/\theta}{1 - \frac{\phi}{x} + 1/\theta} \gamma\lambda,$$

which yields the unique value

$$x_Z = \arg \text{solve} \left\{ \alpha(x - \phi) = \rho - \lambda + \frac{1 + 1/\theta}{1 - \frac{\phi}{x} + 1/\theta} \gamma\lambda \right\} > \phi.$$

As x keeps growing, the variety innovation line (18) enters the positive quadrant and eventually catches up and overtakes the quality innovation line (19), at which point the Nash equilibrium with both vertical and horizontal R&D takes hold. Specifically, given that along this path agents anticipate $z > 0$ at the switch point, (20) yields that $n > 0$ if

$$x > x_N \equiv \frac{(1 - \alpha) \phi - \rho + \lambda - \gamma \lambda}{1 - \alpha - (\rho - \lambda) \pi} > x_Z.$$

To identify the condition on the fundamentals that yields which one of the two scenarios arises it is then sufficient to check for what values of the parameters (18) and (19) go through the origin for the same value of x . According to (19) $z = 0$ for $n = 0$ if

$$\alpha (x - \phi) = \rho - \lambda + \frac{1 + 1/\theta}{1 - \frac{\phi}{x} + 1/\theta} \gamma \lambda$$

while according to (18) $n = 0$ for $z = 0$ if

$$1 - \frac{\phi}{x} = (\rho - \lambda) \pi.$$

These two equations hold for the same x when

$$\alpha \phi \frac{(\rho - \lambda) \pi}{1 - (\rho - \lambda) \pi} = \rho - \lambda + \frac{1 + 1/\theta}{(\rho - \lambda) \pi + 1/\theta} \gamma \lambda,$$

which yields

$$\alpha = \bar{\alpha} \equiv \frac{\rho - \lambda + \frac{1 + 1/\theta}{(\rho - \lambda) \pi + 1/\theta} \gamma \lambda}{\phi \frac{(\rho - \lambda) \pi}{1 - (\rho - \lambda) \pi}}.$$

Now note that the firm size threshold at which the variety innovation line (18) goes through the origin is independent of α while the firm size threshold at which the quality innovation line (19) goes through the origin is decreasing in α . It then follows that for $\alpha < \bar{\alpha}$ the first scenario occurs, while the second occurs for $\alpha > \bar{\alpha}$. Substituting in the expression for $\bar{\alpha}$ derived above yields the inequalities in the text of the proposition.

8.6 Proof of Proposition 5

The definition of firm size (14) and the reduced-form production function (12) yield the differential equation

$$\frac{\dot{x}}{x} = y - n - z = \gamma \lambda - (1 - \sigma) n.$$

The behavior of entrants in (18) and (20) then yields the law of motion of firm size that holds in each case.

- In the variety-first case, I have:

$$\dot{x} = \begin{cases} \gamma \lambda x & x \leq x_N \\ \bar{\nu} (\bar{x}^* - x) & x_N < x \leq x_Z \\ \nu (x^* - x) & x > x_Z \end{cases},$$

where:

$$\begin{aligned}\bar{\nu} &\equiv \frac{1-\sigma}{\pi} \left[1 - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1-\sigma} \right) \right]; \\ \bar{x}^* &\equiv \frac{\phi}{1 - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1-\sigma} \right)}; \\ \nu &\equiv \frac{1-\sigma}{\pi - \sigma/x} \left[1 - \alpha - \pi \left(\rho - \lambda + \frac{\gamma\lambda}{1-\sigma} \right) \right].\end{aligned}$$

Without loss of generality, I approximate $\sigma/x \cong 0$ for $x > \max\{x_N, x_Z\}$ so that the coefficient ν becomes constant and therefore the law of motion of x is piecewise linear. Figure 7 shows the phase diagram for x with the approximation (in red) that yields a linear differential equation in the last phase

Figure 7: The phase diagram for firm size with the approximation (in red) that yields a linear differential equation in the last phase

and time t yields

$$x(t) = x_0 e^{\gamma\lambda t}.$$

Since x grows exponentially, there exists a value T_N such that

$$x(T_N) = x_0 e^{\gamma\lambda T_N} = x_N,$$

which yields (27). Integrating the second line between time T_N and time t yields

$$x(t) = x_N e^{\bar{\nu}(T_N-t)} + \bar{x}^* \left(1 - e^{\bar{\nu}(T_N-t)} \right).$$

For $\bar{x}^* > x_Z$ there exists a finite value T_Z such that

$$x(T_Z) = x_N e^{\bar{\nu}(T_N-T_Z)} + \bar{x}^* \left(1 - e^{\bar{\nu}(T_N-T_Z)} \right) = x_Z,$$

which yields (28). Thereafter the economy follows the third line and converges to the value x^* given in (22). Integrating between time T_Z and time t yields

$$x(t) = x_Z e^{\nu(T_Z-t)} + x^* \left(1 - e^{\nu(T_Z-t)} \right).$$

- In the quality-first case, instead, I have:

$$\dot{x} = \begin{cases} \gamma\lambda x & x \leq x_Z \\ \gamma\lambda x & x_Z < x \leq x_N \\ \nu(x^* - x) & x > x_N \end{cases}.$$

Integrating the first line of between time 0 and time t yields $x(t) = x_0 e^{\gamma\lambda t}$. Accordingly, there exists a value T_Z such that

$$x(T_Z) = x_0 e^{\gamma\lambda T_Z} = x_Z,$$

which yields (30). After crossing this threshold firm size keeps growing exponentially and therefore *must* cross the entry threshold x_N in finite time. Integrating between time T_Z and time t yields

$$x(t) = x_Z e^{\gamma\lambda(t-T_Z)}.$$

There thus exists a value T_N such that

$$x(T_N) = x_Z e^{\gamma\lambda(T_N-T_Z)} = x_N,$$

which yields (31). Thereafter the economy follows the third line of (29), which is identical to the third line of (26), and converges to x^* . Integrating between time T_N and time t yields

$$x(t) = x_N e^{\nu(T_N-t)} + x^* (1 - e^{\nu(T_N-t)}).$$

8.7 The policy functions for the quality-first case

Let $b \equiv C/Y$. Consider $b(x)$ first. Recall that for $x > x_N$ one has $b(x) = b^* \equiv (1 - \theta) [\theta(\rho - \lambda)\pi + 1]$. At issue, then, is only what happens in the region where entrants are not active. Consider first the case of variety innovation first. This is straightforward since firms turn on quality growth only after free entry already applies and thus $b(x)$ is given by (17) evaluated at $z = 0$ over the entire range $\phi \leq x \leq x_N$. The case of quality innovation first is more interesting, since it requires taking into account the dynamic feedbacks through $z > 0$. As before, over the range $\phi \leq x \leq x_Z$ the function $b(x)$ is given by (17) evaluated at $z = 0$. To characterize it over the range $x_Z < x \leq x_N$, substitute $z = \alpha(x - \phi) - \rho + \lambda - \gamma\lambda - (c - y)$ into (17) and rearrange terms to get

$$\frac{\dot{b}}{b} = \frac{x}{\theta} \left(\frac{b}{1-\theta} - 1 \right) - \rho + \lambda - \gamma\lambda - (1 - \alpha)(x - \phi).$$

This yields the $\dot{b} \geq 0$ locus

$$b \geq (1 - \theta) \left[1 + \theta \frac{\rho + \lambda + \gamma\lambda + (1 - \alpha)(x - \phi)}{x} \right].$$

The dynamics then imply that the unique equilibrium trajectory is for the economy to jump on the saddle path in (x, b) space that converges to (x^*, b^*) . Writing

$$\frac{\dot{b}}{\dot{x}} = \frac{db}{dx} = \frac{b \left[\frac{x}{\theta} \left(\frac{b}{1-\theta} - 1 \right) - \rho + \lambda - \gamma\lambda - (1 - \alpha)(x - \phi) \right]}{\gamma\lambda x}$$

characterizes the saddle path more sharply. Although this partial differential equation doesn't have a closed-form solution, it is straightforward to show that the function $b(x)$ has the same derivative from the left and the right at $x = x_Z$ and approaches the value b^* with zero derivative at $x = x_N$:

$$\begin{aligned}\frac{db(x_Z)}{dx} &= \frac{b(x_Z) \left[\frac{x_Z}{\theta} \left(\frac{b(x_Z)}{1-\theta} - 1 \right) - \rho + \lambda - \gamma\lambda - (1-\alpha)(x_Z - \phi) \right]}{\gamma\lambda x_Z} \\ &= \frac{\theta(1-\theta)\phi}{x_Z^2}; \\ \frac{db(x_N)}{dx} &= \frac{b(x_N) \left[\frac{x_N}{\theta} \left(\frac{b(x_N)}{1-\theta} - 1 \right) - \rho + \lambda - \gamma\lambda - (1-\alpha)(x_N - \phi) \right]}{\gamma\lambda x_N} \\ &= 0.\end{aligned}$$

In other words, it is increasing, concave and has no kinks. Solving (17) for z yields

$$z(x) = x - \frac{x}{\theta} \left(\frac{b(x)}{1-\theta} - 1 \right) - \phi.$$

Once again, it is straightforward to show that $z(x)$ starts out at $x = x_Z$ with zero derivative and approaches the line $z(x) = \alpha(x - \phi) - \rho + \lambda - \gamma\lambda$, which holds for $x > x_N$, with positive derivative:

$$\begin{aligned}\frac{dz(x_Z)}{dx} &= 1 - \frac{1}{\theta} \left(\frac{b(x_Z)}{1-\theta} - 1 \right) - \frac{x_Z}{\theta} \frac{db(x_Z)/dx}{1-\theta} \\ &= \frac{\phi}{x_Z} - \frac{x_Z}{\theta} \frac{1}{1-\theta} \frac{\theta(1-\theta)\phi}{x_Z^2} = 0; \\ \frac{dz(x_N)}{dx} &= 1 - \frac{1}{\theta} \left(\frac{b(x_N)}{1-\theta} - 1 \right) - \frac{x}{\theta} \frac{db(x_N)/dx}{1-\theta} = 1 - (\rho - \lambda)\pi > \alpha.\end{aligned}$$

The function $z(x)$ exhibits a kink at $x = x_N$ because when entry begins quality innovation attracts only a fraction of the economy's saving flow, which is now a constant fraction of final output.

References

- [1] Aghion, P., and Howitt, P., 1998, *Endogenous Growth Theory*. Cambridge, MA: MIT University Press.
- [2] Barro, R., and Sala-i-Martin, X., 2004, *Economic Growth* (Second Edition). Cambridge MA: MIT Press.
- [3] Desmet, K. and Parente, S.L., 2012, The evolution of markets and the revolution of industry: a unified theory of growth, *Journal of Economic Growth*, 17: 205-234
- [4] Easterlin, R.A., 1996, *Growth Triumphant: The Twenty-first Century in Perspective*. Ann Arbor: The University of Michigan Press
- [5] Galor, O., 2005, From stagnation to growth: unified growth theory. In P. Aghion and S.N. Durlauf (eds) *Handbook of Economic Growth*, Vol. 1A, pp. 171-293. Elsevier: Amsterdam.

- [6] Galor, O., 2011, *Unified Growth Theory*. Princeton NJ: Princeton University Press.
- [7] Galor, O., and Weil, D.N., 2000, Population, technology and growth: From the Malthusian regime to the demographic transition, *American Economic Review*, 110: 806-828.
- [8] Ha, J., and Howitt, P., 2007, Accounting for trends in productivity and R&D: A Schumpeterian Critique of Semi-Endogenous Growth Theory, *Journal of Money, Credit, and Banking*, 33: 733-74.
- [9] Jones, C.I., 2005, Growth and ideas. In P. Aghion and S.N. Durlauf (eds) *Handbook of Economic Growth*, Vol. 1B, pp. 1063-1111. Elsevier: Amsterdam.
- [10] Jones, E.L., 1988, *Growth Recurring: Economic Change in World History*. Ann Arbor: University of Michigan Press.
- [11] Laincz, C., and Peretto, P.F., 2006, Scale effects in endogenous growth theory: An error of aggregation, not specification, *Journal of Economic Growth*, 11: 263-288.
- [12] Lucas, R., 2002, *Lectures on Economic Growth*. Cambridge MA: Harvard University Press.
- [13] Madsen, J.B., 2008, Semi-endogenous versus Schumpeterian growth models: Testing the knowledge production function using international data, *Journal of Economic Growth*, 13: 1-26.
- [14] Madsen, J.B., 2010. The anatomy of growth in the OECD since 1870: The transformation from the post-Malthusian growth regime to the modern growth epoch, *Journal of Monetary Economics*, 57, 753-767.
- [15] Madsen, J.B., and Ang, J.B., 2011, Can second-generation endogenous growth models explain the productivity trends in the Asian miracle economies?, *Review of Economics and Statistics*, forthcoming.
- [16] Madsen, J.B., Ang, J.B., and Banerjee, R., 2010, Four centuries of British economic growth: the roles of technology and population, *Journal of Economic Growth*, 15: 263-290.
- [17] Mokyr J., 2005, Long-term growth and the history of technology. In P. Aghion and S.N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 1B, pp. 1113-1180. Elsevier: Amsterdam.
- [18] Mokyr J., 2010, The contribution of economic history to the study of innovation and technical change: 1750-1914. In B.H. Hall and N. Rosenberg (eds), *Handbook of The Economics of Innovation*, Vol. 1, pp., 11-50. Elsevier: Amsterdam.
- [19] Mokyr J. and Voth J., 2010, Understanding growth in Europe 1700-1870: Theory and evidence. In S. Broadberry and K.H. O'Rourke (eds), *The Cambridge Economic history of Modern Europe*, Vol. 1: 1700-1870, pp. 7-42. Cambridge: Cambridge University Press.

- [20] Peretto, P. F., 1998, Technological change, market rivalry, and the evolution of the capitalist engine of growth. *Journal of Economic Growth*, 3, 53-80.
- [21] Peretto, P. F. (1999a). Cost reduction, entry, and the interdependence of market structure and economic growth. *Journal of Monetary Economics*, 43, 173-195.
- [22] Peretto, P. F. (1999b). Industrial development, technological change, and long-run growth. *Journal of Development Economics*, 59, 389- 417.
- [23] Peretto, .P.F., 2013, Through scarcity to prosperity: A theory of the transition to sustainable growth, manuscript.
- [24] Simon J.L., 2000, *The Great Breakthrough and its Cause* (edited by T. Kuran). Ann Arbor: The University of Michigan Press.
- [25] Strulik, H., Prettnner, K., Prskawetz, A., 2013, The past and future of knowledge-based growth, manuscript.
- [26] Ulku, H., 2007, R&D, Innovation, and Growth: Evidence from Four Manufacturing Sectors in OECD Countries, *Oxford Economic Papers*, 59: 513-535.