Robust endogenous growth and the nature of medium-frequency cycles.

Tom Holden[1], *School of Economics, University of Surrey*

Abstract: *Existing models of endogenous growth generate implausibly large trend breaks in output when augmented with standard business cycle shocks. This paper presents a model without this deficiency, yet still capable of generating large medium-frequency fluctuations around the trend. Ensuring the robustness of the trend requires that we eliminate the strong scale effects and knife edge assumptions that plague most growth models. In our model, medium-frequency fluctuations arise from changes in the proportion of industries producing patent protected products. However, variations in the number of firms within each industry ensure that process improvement incentives remain roughly constant. An estimated version of the model matches well the observed pattern of medium frequency cycles.*

Keywords: *medium frequency cycles, patent protection, scale effects*

JEL Classification: *E32, E37, L16, O31, O33, O34*

[1] Contact address: School of Economics, FBEL, University of Surrey, Guildford, GU2 7XH, England

Contact email: thomas.holden@gmail.com Web-site: http://www.tholden.org/

The author would like to thank the participants in numerous seminars at which versions of this paper were presented. Further special thanks are due to the following individuals for helpful discussions: Guido Ascari, Eric Bartelsman, Paul Beaudry, Beatriz de Blas, Cristiano Cantore, Diego Comin, Ana-Maria Dumitru, Martin Ellison, Fabio Ghironi, Bart Hobijn, Tim Kehoe, Paul Levine, Simon Wren-Lewis, Omar Licandro, Ioana Moldovan, Peter Neary, Joaquim Oliveira Martins, Christopher Nekarda, Walter Park, Pietro Peretto, Manuel Santos, Diego Restuccia and Klaus Wälde. Financial support provided to the author by the ESRC is also greatly appreciated. Furthermore, the author gratefully acknowledges the use of the Oxford e-Research Centre's Windows Compute Cluster and the University of Surrey FBEL cluster for the numerical computations performed in this paper, and the use of data from the Nuffield College Data Library.

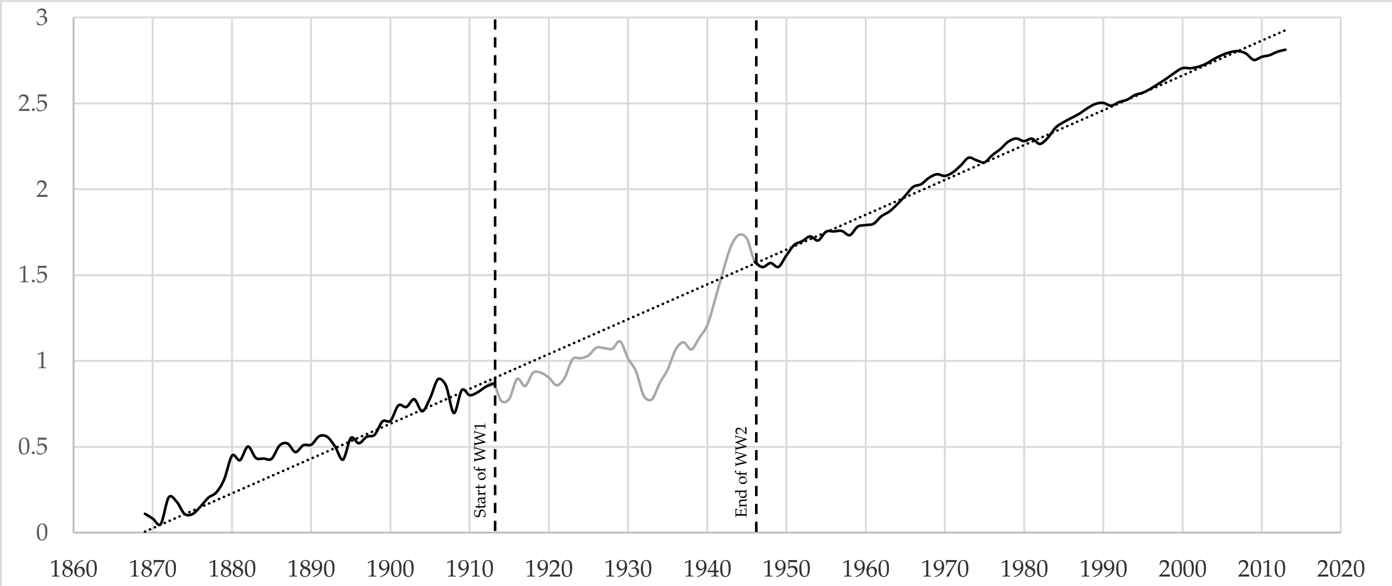

Figure 1: The logarithm of US real GDP per capita,[2] 1869-2013.

The dotted line is a linear trend fitted to the post-war period (1946-2013), which follows the second vertical dashed line.

Figure 1, showing the trend in US real GDP per capita, is surely familiar to everyone in the discipline. Despite this though, it has many features that remain puzzling. As may be seen, a trend fitted to the post-WW2 period also fits well pre-WW1.[3] Indeed, as far back as 1871 real GDP per capita is essentially on the post-WW2 trend. That neither the Great Depression nor two World Wars should have knocked GDP per capita off its log-linear trend is either a spectacular coincidence, or a hint to the deep properties of the mechanisms behind long-run growth at the technological frontier. This is all the more remarkable when considered in light of the structural changes taking place elsewhere in the economy, not least the increase in female labour force participation, and the shift from agriculture to manufacturing to services.

Figure 1 also makes clear the size and persistence of the fluctuations around the long-run growth trend. Even if one puts aside the Great Depression and the war economy years, many of the cycles visible have periods of over ten years, and magnitudes of over 10%. These are not the traditional business cycles. Indeed, these cycles are at frequencies that would be swept into the trend by standard trend/cycle filters. Instead, these are the medium-frequency cycles which Comin and Gertler (2006) did so much to bring to the attention of the profession.

However, traditional models of stochastic endogenous growth struggle to generate large medium-frequency fluctuations that nonetheless eventually return to trend. Indeed, generating robustly log-linear growth is even a challenge in non-stochastic models, with knife-edge restrictions apparently inescapable if counter-factual strong scale effects of population on growth rates are to be removed. In this paper, we produce a novel model of stochastic endogenous growth that solves both problems. It generates both robustly exponential long-run growth, without knife-edge assumptions, and large medium-frequency cycles around that long-run trend.

---

[2] Real GDP, 1869-1928, and population (including armed forces overseas), 1869-1951, are taken from the Historical Statistics of the United States Millennial Edition Online (HSUS), (Carter et al. 2006). Real GDP 1929-2013 is taken from the BEA NIPA tables, with "Intellectual Property Products" (IPP) removed from output via inverting the Fisher Index formula, to bring the data into line with the former practice of treating (IPP) as intermediates. Population 1952-2013 is taken from the FRED series POP, choosing the July value each year in line with the HSUS data. The start date of 1869 was chosen as prior to this point the HSUS data is generated via interpolation.

[3] If an AR(2) + trend model is fitted to this data, allowing for structural breaks in the trend at the marked points, then we cannot reject the null of no breaks at 5%.

In fact, the solutions to both problems are closely related. If a model only generates exponential growth via a knife-edge assumption (including the assumption of precisely exponential population growth), then standard business cycle shocks will inevitably knock the model away from perfectly counter-balancing the scale effect, leading to a permanent break in the trend of the GDP. For example, in models in which the scale effect is removed via new product creation exactly matching population growth, a temporary shock to population will produce a break in trend GDP unless products can be invented implausibly quickly.

Our chief novelty on the endogenous growth side is allowing for a varying number of firms within each industry. Following the recent endogenous growth literature, our model features both product and process innovation, but it is this third margin of industry entry and exit that proves crucial for generating robustly exponential growth. In an important paper, Li (2000) showed that previous models of dual margin endogenous growth do not generate exponential growth once asymmetric spill-overs between product and process innovation are allowed for. We allow for such asymmetric spillovers, but are still able to get asymptotically exponential growth for a positive measure of parameters, thanks to this additional margin.

Within this structure, we are able to generate large medium-frequency fluctuations as follows. In our model, the returns to inventing a new product are higher in a boom due to the higher demand. As a result, during periods of expansion, the rate of creation of new products increases, in line with the evidence of Broda and Weinstein (2010). Due to a first mover advantage, patent protection, or reverse-engineering difficulties, the inventors of these new products will be able to extract rents from them, increasing the costs manufacturing firms face if they wish to produce the new product. These higher costs lead to lower competition in new industries, increasing mark-ups and thus increasing firms' incentives to perform the R&D necessary to catch-up with and surpass the frontier, for basically Schumpeterian reasons. Consequently, the higher proportion of industries that are relatively new in a boom will lead to higher aggregate productivity, lower dispersion of both productivity levels and growth rates, as well as higher mark-ups. Since the length of time for which inventors can extract rents will be determined by the effective duration of patent-protection, this effect will naturally work at medium frequencies. However, since we allow both for the creation of new industries (producing new products) and for varying numbers of firms within each industry, even in the short-run the demand faced by any given firm will be roughly constant, meaning that our model will not produce large deviations from linear growth.

The fact that our endogenous growth model has substantial implications for dynamics increases the range of evidence that we may bring to support it. Formal evidence on the low variance of output per capita at frequency zero was presented by Cochrane (1988), and in TODO we present further evidence that US GDP per capita returns to trend at long lags (at least eight years after the initial shock). Evidence for the pro-cyclicality of TFP has been presented by Bils (1998) and Campbell (1998) amongst others, with Comin and Gertler (2006) showing that the evidence is particularly clear at medium-frequencies. The counter-cyclicality of productivity dispersion has been shown by Kehrig (2011), with evidence on the counter-cyclicality of the dispersion of productivity growth rates provided by e.g. Eisfeldt and Rampini (2006) and Bachmann and Bayer (2009). Evidence for the pro-cyclicality of aggregate mark-ups has been presented by Boulhol (2007) and Nekarda and Ramey (2010). Nekarda and Ramey also show that mark-ups lead output at business-cycle frequencies. In TODO, we present further evidence that this relationship continues to hold at medium-frequencies, with mark-ups being pro-cyclical providing the data is filtered with a cut-off below sixteen years. Boulhol (2007) also shows that although aggregate mark-ups are pro-cyclical, the mark-ups in any particular industry tend to be counter-cyclical. This apparent contradiction is readily explained by our model, as the increase in competition in any particular industry will lead to a decline in mark-ups in that industry (much as in the models of Bilbiie, Ghironi, and Melitz (2012) and Jaimovich (2007)), despite the fact that aggregate mark-ups have increased due to the greater proportion of industries with relatively high mark-ups.

Direct evidence for the importance of our mechanism comes from a number of sources. Balasubramanian and Sivadasan (2011) find that firms holding patents have 17% higher TFP levels on average, and additionally find that firms that go from not holding a patent to holding one experience a 7.4% increase in a fixed effects measure of productivity, suggesting that industries producing patent-protected products are indeed significantly more productive. Serrano (2007) finds that although aggregate patenting is only weakly correlated with aggregate TFP, a measure of the number of patents whose ownership is transferred is strongly related to productivity. He argues that there is a great deal of noise in measures of total patent activity, since so many patents are never seriously commercialised. Patent transfers are usually observed though when their purchaser intends to begin exactly such a commercialisation. Thus, patent transfers provide a proxy for the commencement of production of new patented-products, one that is found to be highly pro-cyclical. Finally, in TODO we present new evidence that longer patent protection significantly increases the share of GDP variance attributable to cycles of medium frequency, suggesting that patent protection plays an important role in the mechanism generating medium frequency cycles in reality.

Previous papers have introduced endogenous productivity improvement into business cycle models (e.g. Comin and Gertler (2006), Comin (2009), Comin, Gertler, and Santacreu (2009), Phillips and Wrase (2006), Nuño (2008; 2009; 2011)), or looked at cycles in growth models (e.g. Bental and Peled (1996), Matsuyama (1999), Wälde (2005), Francois and Lloyd-Ellis (2008; 2009), Comin and Mulani (2009)). However, all of these papers have problems with scale effects, either in the long run, or in the short run following standard shocks, and thus all of them would predict counter-factually large trend breaks in output in the presence of standard business cycle shocks. Furthermore, it is not obvious how these scale effects could be removed without destroying the papers' mechanisms for generating aggregate TFP movements. For example, the papers of Wälde (2005) and Phillips and Wrase (2006) rely on there being a small finite number of sectors. Removing the scale effect would mean allowing this number to grow over time with population, meaning the variance of productivity would rapidly go to zero. Indeed, this happens endogenously in the model of Horii (2011). Many models of endogenous mark-up determination (e.g. Bilbiie, Ghironi, and Melitz (2012) or Jaimovich (2007)) have a similar problem, with the presence of a small finite number of industries being crucial for explaining the observed variance of mark-ups. Indeed, Bilbiie, Ghironi, and Melitz (2011) write that "reconciling an endogenous time-varying markup with stylized growth facts (that imply constant markups and profit shares in the long run) is a challenge to growth theory". By disentangling the margins of firm entry and product creation, we will be able to answer this challenge.

The paper of most relevance to our work is that of Comin and Gertler (2006), who made the important contribution of bringing the significance of medium-frequency cycles to the attention of the profession. Additionally, their theoretical model, like ours, stresses the links between mark-up variation and productivity growth. Unfortunately, however, it is a model with strong scale effects removed via knife-edge assumptions,[4] with the inevitable consequence that the driving mark-up shock produces a counter-factual trend break in productivity. Furthermore, beyond this trend break, the model generates little endogenous persistence, and also counter-factually predict that increases in mark-ups lead to falls in output, contrary to the empirical evidence of Nekarda and Ramey (2010).[5] While there is room to disagree with the recent empirical work on the cyclicality of output, the fact that mark-up increases lead output increases is much more robustly established.[6] We conclude that the literature still lacks a model of productivity capable of explaining both its short run and its long run behaviour.

In this paper, we present a model capable of doing exactly this. In order to robustly remove scale effects, as discussed above it will feature a varying number of industries, each of which will contain a varying number of firms. We do not wish to make any exogenous assumptions on the differences between industries producing patented products versus those producing unpatented ones, so in order to match the medium-frequency behaviour of productivity and mark-ups it is important that our model allow endogenous variation in these quantities across industries. Were we to assume free transfer of technologies across industries there would be too little difference in productivity between patent-protected and un-patent-protected industries, and hence we would not be able to generate medium-frequency cycles. Equally, were we to assume technology transfer across industries was impossible, then it would be legitimate to inquire whether the difference between these industry types was implausibly large, as perhaps firms in non-protected industries would find it optimal to perform technology transfer even if they did not find it optimal to perform any research. Consequently, in modelling the endogenous productivity in each industry we will allow firms both to perform research, and to perform a costly process of catch-up to the frontier we shall term appropriation.

To make clear the strength of the amplification and persistence mechanism presented here, we initially omit capital from the model, and focus only on the impulse responses to non-persistent shocks when we discuss our model's qualitative behaviour in section TODO. We then go on to embed our mechanism within a richer TODO RBC model, which we proceed to empirically test via structural estimation.

1. Reduced form evidence

TODO Coincidence? Not true for other countries? Not true when you divide by number of workers. Bootstrap test?

TODO Reincorporate old reduced form evidence

2. The core model

We begin by describing the core of both the simple and the richer models. This takes the form of a sector producing perishable "widgets". In the simple model, these widgets will be the sole input to the production of the final good, whereas in the richer model, these widgets will be combined with labour and capital. For simplicity, the production of widgets will not itself require capital in either model. Although our use of this

[4] The social value of the aggregate capital stock enters in multiple places without exponent, in order to capture the idea that "operating costs are proportional to the sophistication of the economy". Had (say) frontier technology been used in its place, the generated impulse responses would likely have been quite different, and had it entered with a non-unit exponent then the model would not have possessed a balanced growth path.

[5] Care must be taken to match measures of the aggregate mark-up. If we measure the aggregate mark-up by the inverse labour share, then holding labour supply and output constant, an increase in wage mark-ups decreases aggregate mark-ups. However, at reasonable calibrations of the Comin and Gertler (2006) model, an increase in wage mark-ups results in such a drop in labour supply that the inverse labour share increases.

[6] See TODO for a discussion of the evidence on whether this correlation is causal.

widget sector is chiefly motivated by expositional brevity, the reader may like to think of the widgets as high-tech goods such as computer chips.

The widget production sector comprises a continuum of narrow industries, each of which contains finitely many firms producing a unique product. The measure of industries is increased by the invention of new products, which start their life patent-protected. However, we assume that product inventors lack the necessary human capital to produce their product at scale themselves, and so they must licence out their patent to manufacturing firms. The duration of patent-protection is given by a geometric distribution, in line with Serrano's (2010) evidence on the large proportion of patents that are allowed to expire early, perhaps because they are challenged in court or perhaps because another new product is a close substitute. An earlier working-paper version of this model (Holden 2011) considered the fixed duration case, which is somewhat less tractable. Allowing for a distribution of protection lengths also allows us to give a broader interpretation to protection within our model. Even in the absence of patent protection, the combination of contractual agreements such as NDAs, and difficulties in reverse engineering, is likely to enable the inventor of a new product to extract rents for a period.

Our model of endogenous competition within each industry is derived from Jaimovich (2007). We chose the Jaimovich model as it is a small departure from the standard Dixit-Stiglitz (1977) set-up, and leads to some particularly neat expressions. Similar results could be attained with Cournot competition, or the translog form advocated by Bilbiie, Ghironi, and Melitz (2012). One important departure from the Jaimovich model is that in our model, entry decisions take place one period in advance. This is natural as we wish to model research as taking place after entry but before production. Productivity within a firm is increased by performing research or appropriation. We regard process research as incremental, with regular small changes rather than the unpredictable jumps found in Schumpetarian models (Aghion and Howitt 1992; Wälde 2005; Phillips and Wrase 2006).

Throughout, we assume that only products are patentable,[7] and so by exerting effort firms are able to "appropriate" process innovations from other industries to aid in the production of their own product. This appropriation is costly since technologies for producing other products will not be directly applicable to producing a firm's own product. We assume that technology transfer within an industry is costless however, due to intra-industry labour flows and the fact that all firms in an industry are producing the same product. This is important for preserving the tractability of the model, as it means that without loss of generality we may think of all firms as just existing for two periods, in the first of which they enter and perform research, and in the second of which they produce.

The broad timing of our model is as follows. At the beginning of period t invention takes place, creating new industries. All holders of current patents (including these new inventors) then decide what level of licence fee to charge. Then, based on these licence fees and the level of overhead costs, firms choose whether to enter each industry. Next, firms perform appropriation, raising their next-period productivity towards that of the frontier, then research, further improving their productivity next period. In period $t + 1$, they then

[7] This is broadly in line with the law in most developed countries: ideas that are not embedded in a product (i.e. a machine) generally have at most limited patentability. In the U.S., the most recent Supreme court decision found that the following was "a useful and important clue" to the patentability of processes (*Bilski v. Kappos*, 561 U.S. ___ (2010)): "a method claim is surely patentable subject matter if (1) it is tied to a particular machine or apparatus, or (2) it transforms a particular article into a different state or thing" (*In re Bilski*, 545 F.3d 943, 88 U.S.P.Q.2d 1385 (Fed. Cir. 2008)). This "machine or transformation" test was widely believed at the time to have ended the patentability of business processes (The Associated Press 2008), and this position was only slightly softened by *Bilski v. Kappos*.

produce using their newly improved production process. Meanwhile, a new batch of firms will be starting this cycle again.

We now give the detailed structure of the widget production sector.

2.1. Aggregators

The aggregate "widget" good is produced by a perfectly competitive industry from the aggregated output $X_t(i)$ of each industry $i \in \mathbb{I}_{t-1} \subset \mathbb{R}$, using the following Dixit-Stiglitz-Ethier (Dixit and Stiglitz 1977; Ethier 1982) style technology:

$$X_t = |\mathbb{I}_{t-1}| \left[\frac{1}{|\mathbb{I}_{t-1}|} \int_{i \in \mathbb{I}_{t-1}} X_t(i)^{\frac{1}{1+\lambda}} di \right]^{1+\lambda}$$

where $\frac{1+\lambda}{\lambda}$ is the elasticity of substitution between goods and where the exponent on the measure of industries ($|\mathbb{I}_{t-1}|$)[8] has been chosen to remove any preference for variety in consumption.[9] We normalize the price of the aggregate widget good to 1.

Similarly, each industry aggregate good $X_t(i)$ is produced by a perfectly competitive industry from the intermediate goods $X_t(i,j)$ for $j \in \{1, \dots, J_{t-1}(i)\}$,[10] using the technology:

$$X_t(i) = J_{t-1}(i) \left[\frac{1}{J_{t-1}(i)} \sum_{j=1}^{J_{t-1}(i)} X_t(i,j)^{\frac{1}{1+\eta\lambda}} \right]^{1+\eta\lambda}$$

where $\eta \in (0,1)$ controls the degree of differentiation between firms, relative to that between industries.

2.2. Intermediate firms

2.2.1. Pricing

Firm j in industry i has access to the linear production technology $X_t(i,j) = A_t(i,j) L_t^{\mathrm{P}}(i,j)$ for production in period t, where $A_t(i,j)$ is their productivity and $L_t^{\mathrm{P}}(i,j)$ is their labour input. As in Jaimovich (2007), strategic profit maximisation then implies that in a symmetric equilibrium, the price of the good in industry i is given by $P_t(i) = \left(1 + \mu_{t-1}(i)\right) \frac{W_t^{\mathrm{W}}}{A_t(i,j)} = \left(1 + \mu_{t-1}(i)\right) \frac{W_t^{\mathrm{W}}}{A_t(i)}$, where $\mu_t(i) := \lambda \frac{\eta J_t(i)}{J_t(i) - (1-\eta)} \in (\eta\lambda, \lambda]$ is the industry i mark-up in period $t+1$, W_t^{W} is the wage in the widget sector and $A_t(i) = A_t(i,j)$ is the productivity shared by all firms in industry i in symmetric equilibrium. From aggregating across industries, we then have that $W_t^{\mathrm{W}} = \frac{A_t}{1+\mu_{t-1}}$ where $\frac{1}{1+\mu_t} = \left[\frac{1}{|\mathbb{I}_t|} \int_{i \in \mathbb{I}_t} \left[\frac{1}{1+\mu_t(i)} \right]^{\frac{1}{\lambda}} di \right]^{\lambda}$ determines the aggregate mark-up μ_{t-1} and where:

$$A_t := \frac{\left[\frac{1}{|\mathbb{I}_{t-1}|} \int_{i \in \mathbb{I}_{t-1}} \left[\frac{A_t(i)}{1+\mu_{t-1}(i)} \right]^{\frac{1}{\lambda}} di \right]^{\lambda}}{\left[\frac{1}{|\mathbb{I}_{t-1}|} \int_{i \in \mathbb{I}_{t-1}} \left[\frac{1}{1+\mu_{t-1}(i)} \right]^{\frac{1}{\lambda}} di \right]^{\lambda}}$$

is a measure of the aggregate productivity level.[11]

---

[8] The $t-1$ subscript here reflects the fact that industries are invented one period before their product is available to consumers.

[9] Incorporating a preference for variety would not change the long-run stability of our model. However, it would introduce counterfactual cointegration between productivity and population, suggesting that in reality the preference for variety is at most very weak. (An OLS regression of annual real GDP per capita from 1840 on population and a trend gives a negative coefficient on population, and reduced rank regression of the bivariate cointegrated VAR gives a positive but insignificant coefficient on population in the cointegrating vector.)

[10] Again, the $t-1$ subscript reflects the fact that firms enter one period before production.

[11] Due to the non-linear aggregation, it will not generically be the case that aggregate output is aggregate labour input times A_t. However, the aggregation chosen here is the unique one under which aggregate mark-ups are known one period in advance, as industry mark-ups are.

2.2.2. Sunk costs: rents, appropriation and research

Following Jaimovich (2007), we assume that the number of firms in an industry is pinned down by the zero profit condition that equates pre-production costs to production period profits. Firms raise equity in order to cover these upfront costs, which come from four sources.

Firstly, firms must pay a fixed operating cost of L^{F} units of labour that covers things such as bureaucracy, human resources, facility maintenance, training, advertising, shop set-up and capital installation/creation. Asymptotically, the level of fixed costs will not matter, but including it here will help in our explanation of the importance of patent protection for long run growth.

Secondly, if the product produced by industry i is currently patent-protected, then firms must pay a rent of $\mathcal{R}_t(i)$ units of the consumption good to the patent-holder for the right to produce in their industry. Since all other sunk costs are paid to labour, for convenience we define $L_t^{\mathcal{R}}(i) := \frac{\mathcal{R}_t(i)}{W_t^{\mathrm{W}}}$, i.e. the labour amount equivalent in cost to the rent.

Thirdly, firms will expand labour effort on appropriating the previous process innovations of the leading industry. We define the level of the leading technology within industry i by $A_t^*(i) := \max_{j\in\{1,\dots,J_{t-1}(i)\}} A_t(i,j)$ and the level of the best technology anywhere by $A_t^* := \sup_{i\in[0,\mathbb{I}_{t-1}]} A_t^*(i)$. Due to free in-industry transfer, even without exerting any appropriation effort, firms in industry i may start their research from $A_t^*(i)$ in period t. By employing appropriation workers, a firm may raise this level towards A_t^*.

We write $A_t^{**}(i,j)$ for the base from which firm $j \in \{1,\dots,J_t(i)\}$ will start research in period t. This base is given by the output of the appropriation process, the returns of which we assume to take the form:

$$A_t^{**}(i,j) = \left[A_t^*(i)^\tau + \left(A_t^{*\tau} - A_t^*(i)^\tau\right)\frac{\mathcal{L}_t^{\mathrm{A}}(i,j)}{1+\mathcal{L}_t^{\mathrm{A}}(i,j)}\right]^{\frac{1}{\tau}}, \tag{1.1}$$

where $\tau > 0$ controls whether the catch-up amount is a proportion of the technology difference in levels ($\tau = 1$), log-levels ($\tau = 0$) or anything in between or beyond, and where $\mathcal{L}_t^{\mathrm{A}}(i,j)$ is the effective input to appropriation. This in turn is given by $\mathcal{L}_t^{\mathrm{A}}(i,j) := E_t^{\mathrm{A}}(i)L_t^{\mathrm{A}}(i,j)$, where $L_t^{\mathrm{A}}(i,j)$ is the quantity of labour that the firm devotes to appropriation in period t, and $E_t^{\mathrm{A}}(i) := A_t^*(i)^{-\zeta^{\mathrm{A1}}} A_t^{*\zeta^{\mathrm{A2}}} |\mathbb{I}_t|^{\phi^{\mathrm{A}}} \Psi^{\mathrm{A}}$, where Ψ^{A} is the productivity of appropriation labour, $\zeta^{\mathrm{A1}} > 0$ controls the extent to which appropriation is getting harder over time due to the increased complexity of later technologies, $\zeta^{\mathrm{A2}} \geq 0$ gives the strength of the spillover from process innovation to appropriation, and where $\phi^{\mathrm{A}} \geq 0$ gives the strength of the spillover from product innovation to process appropriation.

This specification captures the key idea that the further a firm is behind the frontier, the more productive will be appropriation. Allowing for spillovers from product and process innovation is essential to demonstrate that in our model endogenous growth is not a knife-edge result, unlike in that of Li (2000). Finally, allowing for appropriation (and research, and invention) to get harder over time is both realistic, and essential for the tractability of our model, since it will lead our model to have a finite dimensional state vector asymptotically, despite all the heterogeneity across industries. In order for this to be the case, we assume that ϕ^{A} is small enough that $A_t^{*-\zeta^{\mathrm{A}}} \mathbb{I}_t^{\phi^{\mathrm{A}}} \to 0$ as $t \to \infty$, where $\zeta^{\mathrm{A}} = \zeta^{\mathrm{A1}} - \zeta^{\mathrm{A2}} > 0$. We will derive sufficient conditions on structural parameters for this in section 3.1.

Fourthly, firms will employ labour in research. If firm $j \in \{1, \dots, J_t(i)\}$ has an effective research input of $\mathcal{L}_t^{\mathrm{R}}(i,j)$, then we assume that its productivity level in period $t+1$ will be given by:

$$A_{t+1}(i,j) = A_t^{**}(i,j)\left(1 + \gamma Z_{t+1}(i,j)\mathcal{L}_t^{\mathrm{R}}(i,j)\right)^{\frac{1}{\gamma}}, \tag{1.2}$$

where $\gamma > 0$ controls the "parallelizability" of research and $Z_{t+1}(i,j) > 0$ is a stationary stochastic process representing the luck component of research, with $\mathbb{E}Z_t \equiv 1$.[12] If $\gamma = 1$, research may be perfectly parallelized, so arbitrarily large quantities may be performed within a given period without loss of productivity, but if γ is large, then, in line with the evidence of Siliverstovs and Kancs (2012), the returns to research decline as the firm attempts to pack more into one period. Much as with appropriation, the effective input to research is given by $\mathcal{L}_t^{\mathrm{R}}(i,j) \coloneqq E_t^{\mathrm{R}}(i)L_t^{\mathrm{R}}(i,j)$, where $L_t^{\mathrm{R}}(i,j)$ is the quantity of labour the firm employs in research in period t, and $E_t^{\mathrm{R}}(i) \coloneqq A_t^{**}(i)^{-\zeta^{\mathrm{R1}}} A_t^{*\zeta^{\mathrm{R2}}} |\mathbb{I}_t|^{\phi^{\mathrm{R}}}\Psi^{\mathrm{R}}$, where Ψ^{R} is the productivity of research labour, ζ^{R1} controls the extent to which research is getting harder over time, $\zeta^{\mathrm{R2}} \geq 0$ gives the spillover from frontier process innovation and $\phi^{\mathrm{R}} \geq 0$ gives the spillover from product innovation to process. We again assume that ϕ^{R} is small enough that $A_t^{*-\zeta^{\mathrm{R}}}|\mathbb{I}_t|^{\phi^{\mathrm{R}}} \to 0$ as $t \to \infty$, where $\zeta^{\mathrm{R}} = \zeta^{\mathrm{R1}} - \zeta^{\mathrm{R2}} > 0$. We also assume that $\zeta^{\mathrm{R1}} > \zeta^{\mathrm{A1}}$, $\zeta^{\mathrm{R2}} \leq \zeta^{\mathrm{A2}}$ and $\phi^{\mathrm{R}} \leq \phi^{\mathrm{A}}$ implying that the difficulty of research is increasing over time faster than the difficulty of appropriation. This is made because research is very much specific to the industry in which it is being conducted, whereas appropriation is a similar task across all industries attempting to appropriate the same technology, and hence is more likely to have been standardised, or to benefit from other positive spillovers.

In the following, for tractability, we will assume that $Z_t(i,j) \coloneqq Z_t$ so that all firms receive the same "idea" shock. We make this assumption chiefly for simplicity, but it may be justified by appeal to common inputs to private research, such as university research output or the availability of new tools, or by appeal to in-period labour market movements carrying ideas with them. We will see in the following that allowing for industry-specific shocks has minimal impact on our results, providing shocks are bounded above.

2.2.3. Research and appropriation effort decisions

Firms are owned by households and so they choose research and appropriation to maximize:

$$\beta\mathbb{E}_t\left[\Xi_{t+1}\left(P_t(i,j) - \frac{W_{t+1}^{\mathrm{W}}}{A_{t+1}(i,j)}\right)X_t(i,j)\right] - \left[L_t^R(i,j) + L_t^A(i,j) + L_t^{\mathbb{R}}(i) + L^F\right]W_t^{\mathrm{W}},$$

where $\beta\Xi_{t+1}$ is the household's stochastic discount factor from period t to $t+1$. It may be shown that, for firms in frontier industries (those for which $A_t^*(i) = A_t^*$), if an equilibrium exists, then it is unique and symmetric within an industry; but we cannot rule out the possibility of asymmetric equilibria more generally.[13] However, since the coordination requirements of asymmetric equilibria render them somewhat

---

[12] Peretto (1999) also looks at research that drives incremental improvements in productivity, and chooses a similar specification. The particular one used here is inspired by Groth, Koch, and Steger (2009).

[13] The equilibrium concept we use is that of pure-strategy subgame-perfect *local* Nash equilibria (SP*L*NE) (i.e. only profitable local deviations are ruled out). We have no reason to believe the equilibrium we find is not in fact a subgame-perfect Nash equilibria (SPNE). Indeed, if there is a pure-strategy symmetric SPNE then it will be identical to the unique pure-strategy symmetric SP*L*NE that we find. Furthermore, our numerical investigations suggest that at least in steady-state, at our calibrated parameters, the equilibrium we describe is indeed an SPNE. (Code available on request.) However, due to the analytic intractability of the second stage pricing game when productivities are asymmetric, we cannot guarantee that it remains an equilibrium away from the steady-state, or for other possible calibrations. However, SPLNE's are independently plausible since they only require firms to know the demand curve they face in the local vicinity of an equilibrium, which reduces the riskiness of the experimentation they must perform to find this demand curve (Bonanno 1988). It is arguable that the coordination required to sustain asymmetric equilibria and the computational demands of mixed strategy equilibria render either of these less plausible than our SPLNE.

implausible, we restrict ourselves to the unique equilibrium in which all firms within an industry choose the same levels of research and appropriation.

Then, providing $\frac{1}{\mu_t(i)} < \min\{\gamma, \tau\}$ and $\gamma > \zeta^{\mathrm{R}1}$ (for the second order conditions[14] and for uniqueness), combining the first order and free entry conditions then gives us that, in the limit as $\operatorname{var} Z_{t+1} \to 0$:[15]

$$\mathcal{L}_t^{\mathrm{R}}(i) = \max\left\{0, \frac{d_t(i) E_t^{\mathrm{R}}(i)\big(L_t^{\mathrm{A}}(i,j) + L_t^{\mathcal{R}}(i) + L_t^{\mathrm{F}}\big) - \mu_t(i)}{\gamma \mu_t(i) - d_t(i)}\right\} \tag{1.3}$$

and:

$$\mathcal{L}_t^{\mathrm{A}}(i) = \max\left\{0, -f_t(i) + \sqrt{\max\{0, f_t(i)^2 + g_t(i)\}}\right\}, \tag{1.4}$$

where $d_t(i) \in (0,1)$[16] is small when firm behaviour is highly distorted by firms' incentives to deviate from choosing the same price as the other firms in their industry, off the equilibrium path (so $d_t(i) \to 1$ as $J_t(i) \to \infty$), and $f_t(i)$ and $g_t(i)$ are increasing in an industry's distance from the frontier,[17] as the further behind a firm is, the greater are the returns to appropriation.

Equations (1.3) and (1.4) mean that research and appropriation levels are increasing in the other sunk costs a firm must pay prior to production, but decreasing in mark-ups. They also mean that the strategic distortions caused by there being a small number of firms within an industry tend to reduce research and appropriation levels. Other sunk costs matter for research levels because when other sunk costs are high, entry into the industry is lower, meaning that each firm receives a greater slice of production-period profits, and so has correspondingly amplified research incentives.

Why mark-up increases decrease research incentives is clearest when those mark-up increases are driven by exogenous decreases in the elasticity of substitution. When products are close substitutes, then by performing research (and cutting its price) a firm may significantly expand its market-share, something that will not happen when the firm's good is a poor substitute for its rivals. When $d_t(i) \approx 1$ (i.e. there are a lot of firms in the industry) firms act as if they faced an exogenous elasticity of substitution $\frac{1+\mu_t(i)}{\mu_t(i)}$, and so when mark-ups are high they will want to perform little research. When $d_t(i)$ is small (i.e. there are only a few firms[18]) then firms' behaviour is distorted by strategic considerations. Each firm realises that if they perform extra research today then their competitors will accept lower mark-ups the next period. This reduces the extent to which research allows market-share expansion, depressing research incentives.

The key thing to note about (1.3) and (1.4) is that research and appropriation are independent of the level of demand, except insomuch as demand affects mark-ups or the level of strategic distortion. This is because

[14] The second order condition for research may be derived most readily by noting that when $d_t(i) \to 1$, (i.e. $J_t(i) \to \infty$) the first order condition for research is identical to the one that would have been derived had there been a continuum of firms in each industry with exogenous elasticity of substitution $\frac{1+\mu_t(i)}{\mu_t(i)}$. That it holds more generally follows by continuity and uniqueness. Since $A_t^{**}(i,j)$ is bounded above, no matter how much appropriation is performed the highest solution of the appropriation first order condition must be at least a local maximum.

[15] The first order and zero profit conditions are reported in the online appendix, section 7.1, where we also derive these solutions. We do not assume $\sigma_Z = 0$ when simulating, but it leads here to expressions that are easier to interpret.

[16] $d_t(i) := 1 - \frac{\omega_t(i)}{1+\omega_t(i)} \frac{(\lambda - \mu_t(i))(\mu_t(i) - \eta\lambda)}{\lambda(1-\eta)\mu_t(i)}$, where $\omega_t(i) := \frac{J_t(i)(1-\eta)}{(J_t(i)-(1-\eta))^2(1+\mu_t(i))}$.

[17] $f_t(i) := \frac{1}{2}\left[1 + \frac{d_t(i)}{\tau\mu_t(i)} \frac{1+(\gamma-\zeta^{\mathrm{R}})\mathcal{L}_t^{\mathrm{R}}(i)}{1+\gamma\mathcal{L}_t^{\mathrm{R}}(i)}\right]\left[1 - \left(\frac{A_t^*(i)}{A_t^*}\right)^{\tau}\right] - 1$, $g_t(i) := \frac{d_t(i)}{\tau\mu_t(i)} \frac{1+(\gamma-\zeta^{\mathrm{R}})\mathcal{L}_t^{\mathrm{R}}(i)}{1+\gamma\mathcal{L}_t^{\mathrm{R}}(i)} E_t^{\mathrm{A}}(i)\left[L_t^{\mathrm{R}}(i) + L_t^{\mathcal{R}}(i) + L^{\mathrm{F}}\right]\left[1 - \left(\frac{A_t^*(i)}{A_t^*}\right)^{\tau}\right] - \left(\frac{A_t^*(i)}{A_t^*}\right)^{\tau}$.

[18] The minimum value of $d_t(i)$ occurs when there is more than one firm in the industry. If there is a single firm in an industry, then, as you would expect, very little research will be performed (because the firm's only incentive to cut prices comes from competition from other industries, competition which is very weak, since those industries are producing poor substitutes to its own good). However, this drop in research incentives works entirely through the mark-up channel, and $d_t(i) \to 1$ as $J_t(i) \to 1$. One intuition for this is that there can be no strategic behaviour when there is only a single firm.

when demand is high there is greater entry, so each firm still faces roughly the same demand. This is essential for removing the short-run scale effect.

In industries that are no longer patent-protected, rents will be zero (i.e. $L_t^{\mathcal{R}}(i) \equiv 0$). Since research is getting harder at a faster rate than appropriation ($\zeta^{\mathrm{R1}} > \zeta^{\mathrm{A1}}$, $\zeta^{\mathrm{R2}} \le \zeta^{\mathrm{A2}}$, $\phi^{\mathrm{R}} \le \phi^{\mathrm{A}}$), at least asymptotically, no research will be performed in these industries. This is because $d_t(i)E_t^{\mathrm{R}}(i)\left[L_t^{\mathrm{A}}(i) + L^{\mathrm{F}}\right] - \mu_t(i)$ is asymptotically negative as $\mu_t(i) \in (\eta\lambda, \lambda]$. For growth to continue forever in the absence of patent protection, we would require that the overhead cost (L^{F}) was growing over time at exactly the right rate to offset the increasing difficulty of research. This does not seem particularly plausible. However, it will turn out that optimal patent rents grow at exactly this rate, so with patent protection we will be able to sustain long run growth even when overhead costs are asymptotically dominated by the costs of research. In the presence of sufficiently severe financial frictions of the "pledgibility constraint" form (Hart and Moore 1994), it may be shown that long run growth is sustainable even without patent protection. We leave the details of this for future work.

2.3. Inventors and patent protection

Each new industry is controlled by an inventor who owns the patent rights to the product the industry produces. Until the inventor's product goes on sale, they can successfully protect their revenue stream through contractual arrangements, such as non-disclosure agreements. This means that even in the absence of patent-protection, an inventor will receive one period of revenues. They then have a probability of $1 - q$ of being granted a patent to enable them to extract rents for a second period. After this, if they have a patent at t, then they face a constant probability of $1 - q$ of having a patent at $t + 1$.

We suppose, however, that inventors lack the necessary human capital to produce their product at scale themselves. Instead, in the period the product is invented, and in each subsequent period in which they have a patent, the inventor optimally chooses the rent $\mathcal{R}_t(i)$ (or equivalently $L_t^{\mathcal{R}}(i)$) to charge all the firms that wish to produce their product.

The reader should have a firm such as Apple in mind when thinking about these inventors. Apple has no manufacturing plants and instead maintains its profits by product innovation and tough bargaining with suppliers.

2.3.1. Optimal rent decisions

Inventors' businesses are also owned by households; hence, an inventor's problem is to choose $L_{t+s}^{\mathcal{R}}(i)$ for $s \in \mathbb{N}$ to maximise their expected present value, subject to an enforceability constraint on rents. Their period t present value is given by:

$$V_t^{\mathbb{I}}(i) := \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s (1-q)^s \left[\prod_{k=1}^{s} \Xi_{t+k}(1 - \delta_{\mathbb{I},t+k})\right] L_{t+s}^{\mathcal{R}}(i) W_{t+s}^{\mathrm{W}} J_{t+s}(i),$$

where $\delta_{\mathbb{I},t}$ is the rate at which products drop out of the widget aggregator, due to changing tastes, or the invention of replacement products.[19] In line with this intuition, we assume that $\delta_{\mathbb{I},t} = \delta_{\mathbb{I}}\left(\frac{|\mathbb{I}_t|}{G_{\mathbb{I}}|\mathbb{I}_{t-1}|}\right)^{\psi} \tilde{\delta}_{\mathbb{I},t}$, where $\psi > 0$ gives the elasticity of product depreciation with respect to product growth, $G_{\mathbb{I}}$ is the value of $\lim_{t\to\infty} \frac{|\mathbb{I}_t|}{|\mathbb{I}_{t-1}|}$ in the non-stochastic limit of the model, and $\tilde{\delta}_{\mathbb{I},t}$ is a stationary stochastic process, with $\mathbb{E}\tilde{\delta}_{\mathbb{I},t} \equiv 1$. We also

[19] This means that were it not for the invention of new products, there would be a measure $(1 - \delta_{I,t})|I_{t-1}|$ of industries producing in period $t + 1$.

assume that the random event of losing patent protection is independent of that of product depreciation, so some (but not all) products will drop out of the aggregator without ever going out of patent protection.

If the rents charged by a patent-holder go too high, a firm is likely to ignore them completely in the hope that either they will be lucky, and escape having their profits confiscated from them by the courts (since proving patent infringement is often difficult), or that the courts will award damages less than the licence fee. This is plausible since the relevant U.S. statute states that "*upon finding for the claimant the court shall award the claimant damages adequate to compensate for the infringement but in no event less than a reasonable royalty for the use made of the invention by the infringer, together with interest and costs as fixed by the court*".[20,21] The established legal definition of a "reasonable royalty" is set at the outcome of a hypothetical bargaining process that took place immediately before production,[22] so patent-holders may just as well undertake precisely this bargaining process before production begins.[23]

This leads patent-holders to set:

$$L_t^{\mathcal{R}}(i) = \frac{1-\rho}{\rho}\left[L_t^{\mathrm{R}}(i) + L_t^{\mathrm{A}}(i) + L^{\mathrm{F}}\right], \tag{1.5}$$

at least for sufficiently large t, where $\rho \in (0,1)$ is the bargaining power of the firm, in the sense of the generalized Nash bargaining solution. The simple form of this expression comes from the fact that a firm's production period revenues (which is what is being bargained over) are precisely equal to the costs they face prior to production, thanks to the free entry condition. A full description of the legally motivated bargaining process is contained in the online appendix, section 7.3, along with a discussion of some technical complications pertaining to off equilibrium play.

From combining (1.3) and (1.5) then, at least for sufficiently large t, in the limit as $\operatorname{var} Z_{t+1} \to 0$, we have that:

$$\mathcal{L}_t^{\mathrm{R}}(i) = \frac{\rho\mu_t(i) - d_t(i)E_t^{\mathrm{R}}(i)\left(L_t^{\mathrm{A}}(i) + L^{\mathrm{F}}\right)}{d_t(i) - \gamma\rho\mu_t(i)}. \tag{1.6}$$

For there to be growth in the long run then, we now just require that $d_t(i) > \gamma\rho\mu_t(i)$, which together with the second order and appropriation uniqueness conditions implies $\gamma\rho < \frac{1}{\mu_t(i)} < \min\{\gamma, \tau\}$.[24] We see that, once optimal rents are allowed for, research is no longer decreasing in mark-ups within an industry, at least for firms at the frontier. Mathematically, this is because the patent-holder sets rents as such a steeply increasing function of research levels. More intuitively, you may think of the patent-holder as effectively controlling how much research is performed by firms in their industry, and as taking most of the rewards from this research. It is then unsurprising that we reach these Schumpeterian conclusions.[25]

---

[20] 35 U.S.C. § 284 Damages.

[21] The reasonable royalty condition is indeed the relevant one for us since our assumption that the patent-holder lacks the necessary human capital to produce at scale themself means it would be legally debatable if they had truly "lost profits" following an infringement (Pincus 1991).

[22] See the online appendix, section 7.3, for evidence supporting this interpretation.

[23] In any case, if we allow for idiosyncratic "idea shocks" firms will wish to delay bargaining until this point anyway, since with a bad shock they will be less inclined to accept high rents. Patent-holders also wish to delay till this point because the more sunk costs the firms have already expended before bargaining begins, the greater the size of the "pie" they are bargaining over.

[24] If the number of firms in protected industries is growing over time then $d_t(i) \to 1$, so asymptotically these conditions are equivalent.

[25] The empirical evidence (Scott 1984; Levin, Cohen, and Mowery 1985; Aghion et al. 2005; Tingvall and Poldahl 2006) suggests that the cross-industry relationship between competition and research takes the form of an inverted-U. Based on the fact that strategic distortions are maximised (i.e. $d_t(i)$ is minimised) when there is a small finite number of firms, one might perhaps hope that this holds in our model too. Unfortunately, the maximum of $\frac{\mu_t(i)}{d_t(i)}$ (and hence of research) as a function of $J_t(i)$ may be shown to always occur at some $J_t(i) < 1$. While fractional entry may be a legitimate way of

2.3.2. Invention

We consider invention as a costly process undertaken by inventors until the expected returns from inventing a new product fall to zero. New products appear at the end of the product spectrum, and are contained in the set of products aggregated into the widget good until they are hit with the product depreciation shock. Therefore, the product index *i* will always refer to the same product, once it has been invented.

There is, however, no reason to think that newly invented products will start with a competitive production process. A newly invented product may be thought of as akin to a prototype: yes, identical prototypes could be produced by the same method, but doing this is highly unlikely to be commercially viable. Instead, there will be rapid investment in improving the product's production process until it may be produced as efficiently as its rivals can be. In our model, this investment in the production process is performed not by the inventor but by the manufacturers. Prototyping technology has certainly improved over time;[26] in light of this, we assume that a new product *i* is invented with a production process of level $A_t^*(i) = S_t A_t^*$, where $S_t \in (0,1)$ is an exogenous process controlling initial relative productivity. We assume that either S_t is stationary, so prototyping technology is keeping up with frontier productivity, or that S_t has a negative growth rate asymptotically, so firms start progressively further behind over time.

Just as we expect process research to be getting harder over time, as all the obvious process innovations have already been discovered, so too we may expect product invention to be getting harder over time, as all the obvious products have already been invented. In addition, the necessity of actually finding a way to produce a prototype will result in the cost of product invention increasing in $A_t^*(i)$, the initial productivity level of the process for producing the new product. Finally, following Li (2000), we allow for spillovers from process innovation to product innovation. As a result of these considerations, we assume that the labour cost is given by $\frac{\mathcal{L}_t^{\mathrm{I}}}{E_t^{\mathrm{I}}}$, where $\mathcal{L}_t^{\mathrm{I}}$ is a stationary stochastic process determining the difficulty of invention, with $\Pr(\mathcal{L}_t^{\mathrm{I}} > \underline{\mathcal{L}}^{\mathrm{I}}) \equiv 1$ and $\mathbb{E}\mathcal{L}_t^{\mathrm{I}} \equiv \mathcal{L}^{\mathrm{I}} > \underline{\mathcal{L}}^{\mathrm{I}}$, and where $E_t^{\mathrm{I}} := (S_t A_t^*)^{-\zeta^{\mathrm{I1}}} A_t^{*\zeta^{\mathrm{I2}}} |\mathbb{I}_{t-1}|^{-\chi}$, with $\zeta^{\mathrm{I2}} \geq 0$ giving the strength of spillovers from process innovation and $\chi \geq 0$ and $\zeta^{\mathrm{I1}} > \zeta^{\mathrm{I2}}$ controlling the rate at which inventing a new product gets more difficult because of, respectively, an increased number of existing products or an increased level of initial productivity.

We are assuming there is free entry of new inventions, so the marginal entrant must not make a positive profit from entering. That is, $|\mathbb{I}_t| \geq (1 - \delta_{\mathbb{I},t})|\mathbb{I}_{t-1}|$ must be as small as possible such that:

$$\frac{\mathcal{L}_t^{\mathrm{I}}}{E_t^{\mathrm{I}}} W_t^{\mathrm{W}} \geq V_t^{\mathbb{I}}(\sup \mathbb{I}_t).$$

If, after a shock, invention can satisfy this equation with equality, then the measure of firms will not have to adjust significantly. However, if the $|\mathbb{I}_t| \geq (1 - \delta_{\mathbb{I},t})|\mathbb{I}_{t-1}|$ constraint binds, then the measure of firms will have to adjust instead, meaning there may be an asymmetry in the response of mark-ups to certain shocks.

modelling niche products that are never fully commercialised, we prefer to explain the inverted-U in the data with reference to the cross-sectional distribution of industries. New industries will start with a production process behind that of the frontier, and thus firms in them will wish to perform large amounts of appropriation and relatively small amounts of research, since appropriation is a cheaper means of increasing productivity for a firm behind the frontier. In the presence of a luck component to appropriation (not included above, for simplicity) this leads new industries to have the highest degree of productivity dispersion, as older industries remain close to the frontier. As a result of this high productivity dispersion, there will be firms in new industries setting both very high, and very low mark-ups, which, combined with the fact they are performing less research than more mature patent-protected industries, would generate an inverted-U.

[26] Examples of recent technologies that have raised the efficiency of prototype production include 3D printing and computer scripting languages such as Python.

2.3.3. The life cycle of an industry

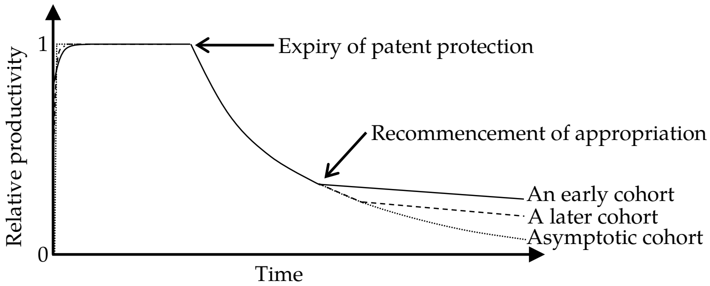

Figure 2: The stylized life cycle of an industry

We are now close to being able to describe the life cycle of an industry in our model. Industries start out with productivity behind that of the frontier, but thanks to the high cost of entry to patent-protected industries, they have strong incentives to catch up to it via appropriation. Once they reach it, thanks to the rents charged by patent holders they will perform research pushing the frontier forward over time. Now, some industries will be hit with the "product-depreciation" shock prior to going out of patent protection, but others will fall out of patent protection first. We have already shown that these industries will not perform research asymptotically, so it just remains to be seen if they will perform appropriation.

Appropriation is performed in an industry i if and only if $g_t(i) > 0$, which, for a non-patent protected industry no longer performing research, is true if and only if:

$$\frac{A_t^*(i)}{A_t^*} < \left(\frac{E_t^{\mathrm{A}}(i)L^{\mathrm{F}}}{E_t^{\mathrm{A}}(i)L^{\mathrm{F}} + \tau\frac{\mu_t(i)}{d_t(i)}}\right)^{\frac{1}{\tau}}.$$

The left hand side of this equation is the relative productivity of the industry compared to the frontier, so industries that are too close to the frontier will not perform appropriation. The right hand side of this equation will be shrinking over time thanks to the increasing difficulty of appropriation, meaning the no-appropriation cut-off point is also declining over time. Indeed, we show in the online appendix, section 7.2, that asymptotically the productivity of non-protected firms is growing at $\left[1+\frac{\zeta^{\mathrm{A2}}}{\tau}\right]\left[1+\frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1} < 1$ times the growth rate of the frontier, plus $\frac{\phi^{\mathrm{A}}}{\tau}\left[1+\frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1}$ times the growth rate of the measure of industries. We go on to show that given our previous assumptions, asymptotically this quantity must be strictly lower than the growth rate of the frontier, and hence the relative productivity of non-protected industries is tending to zero asymptotically. This is empirically plausible since productivity differences across industries have been steadily increasing over time,[27] and it is important for the tractability of our model since it enables us to focus on the asymptotic case in which non-protected firms never perform appropriation. It is also in line with the long delays in the diffusion of technology found by Mansfield (1993) amongst others.

Armed with this knowledge we can depict the full lifecycle of industries from different cohorts. We do this in Figure 2.

[27] Some indirect evidence for this is provided by the increase in wage inequality, documented in e.g. Autor, Katz, and Kearney (2008). Further evidence is provided by the much higher productivity growth rates experienced in manufacturing, compared to those in services (mostly unpatented and unpatentable), documented in e.g. Duarte and Restuccia (2009).

3. Properties of a simple model

We now embed the core model in a stripped down real business cycle model, and investigate its properties. In this model the consumption good will be produced from widgets one for one, and there will be no capital. Thus the market clearing condition in final goods markets takes the form $C_t = X_t$, where as ever, C_t is consumption. As of the 2013 NIPA revision[28], R&D is treated as investment in GDP, so to produce a comparable measure of output we do the same. In our model, total output in units of widgets, Y_t^{W} is equal to expenditure on consumption, research, appropriation and invention, which is:

$$Y_t^{\mathrm{W}} = C_t + \left[\frac{\mathcal{L}_t^{\mathrm{I}}}{E_t^{\mathrm{I}}}(|\mathbb{I}_t| - (1-\delta_{\mathbb{I},t})|\mathbb{I}_{t-1}|) + \int_{i\in\mathbb{I}_t}\left(\frac{\mathcal{L}_t^{\mathrm{R}}(i)}{E_t^{\mathrm{R}}(i)} + \frac{\mathcal{L}_t^{\mathrm{A}}(i)}{E_t^{\mathrm{A}}(i)}\right)J_t(i)\,di\right]W_t^{\mathrm{W}}.$$

To complete the model, we assume there is a unit mass of households, each of which contains N_t members in period t, where N_t is a stochastic process governing population. The representative household chooses consumption C_t and labour supply L_t^{S} to maximise:

$$\mathbb{E}_t \sum_{s=0}^{\infty} \beta^s N_{t+s}\left[\log\frac{C_{t+s}}{N_{t+s}} - \frac{1}{1+\nu}\left(\frac{L_{t+s}^{\mathrm{S}}}{N_{t+s}}\right)^{1+\nu}\right]$$

(where β is the discount rate and ν is the inverse of the Frisch elasticity of labour supply to wages), subject to the aggregate budget constraint that $C_t + B_t = L_t^{\mathrm{S}} W_t^{\mathrm{W}} + B_{t-1}R_{t-1} + \Pi_t$, where B_t is the aggregate number of (zero net supply) bonds bought by households in period t, R_{t-1} is the period t sale price of a (unit cost) bond bought in period $t-1$, and Π_t is the households' period t dividend income. Consequently, the stochastic discount factor is given by $\beta\Xi_t = \beta\frac{N_t C_{t-1}}{N_{t-1}C_t}$.

The model is closed with the labour market clearing condition:

$$L_t^{\mathrm{S}} = \frac{\mathcal{L}_t^{\mathrm{I}}}{E_t^{\mathrm{I}}}[|\mathbb{I}_t| - (1-\delta_{\mathbb{I},t})|\mathbb{I}_{t-1}|] + \int_{i\in\mathbb{I}_t}\left(L_t^{\mathrm{R}}(i) + L_t^{\mathrm{A}}(i) + L^{\mathrm{F}}\right)J_t(i)\,di$$
$$+ W_t^{\mathrm{W}-\frac{1+\lambda}{\lambda}}\frac{X_t}{|\mathbb{I}_{t-1}|}\int_{i\in\mathbb{I}_{t-1}}\frac{1}{A_t(i)}\left(\frac{A_t(i)}{1+\mu_{t-1}(i)}\right)^{\frac{1+\lambda}{\lambda}}di.$$

3.1. Theoretical long run properties

It may be shown from the free entry conditions for firms and inventors that asymptotically, the amount of labour devoted to invention or research is growing at the same rate as the amount of labour devoted to production. Consequently, from the labour first order condition and labour market clearing one, $g_X = g_{A^*} + g_N$, where $g_{\mho}$ is the asymptotic growth rate of the variable $\mho_t$ in the non-stochastic limit. Hence, again from the free entry conditions for firms and inventors:

$$g_{|\mathbb{I}|} = \frac{1}{1+\chi}\left(g_X - g_{W^{\mathrm{W}}} - \zeta^{\mathrm{I}} g_{A^*} - \zeta^{\mathrm{I1}} g_S\right) = \frac{1}{1+\chi}\left(g_N - \zeta^{\mathrm{I}} g_{A^*} - \zeta^{\mathrm{I1}} g_S\right),$$

where $\zeta^{\mathrm{I}} := \zeta^{\mathrm{I1}} - \zeta^{\mathrm{I2}} > 0$. Therefore, if $\chi = \zeta^{\mathrm{I}} = \zeta^{\mathrm{I1}} = 0$ the stock of products will grow at exactly the same rate as population, and away from this special case it may be growing either more slowly or more quickly, since $g_S \le 0$. If invention were to stop asymptotically, eventually there would be no protected industries, and hence no productivity growth. Therefore, for long-run growth, we require that $g_{|\mathbb{I}|} > \log(1-\delta_{\mathbb{I}})$, i.e. $g_N - \zeta^{\mathrm{I1}} g_S - (1+\chi)\log(1-\delta_{\mathbb{I}}) > \zeta^{\mathrm{I}} g_{A^*}$ which will hold providing product depreciation is sufficiently high, regardless of whether or not there is population growth. Even without product depreciation, productivity

[28] http://www.bea.gov/scb/pdf/2013/09%20September/0913_comprehensive_nipa_revision.pdf

growth may be sustained indefinitely in the presence of a declining population if the government offers infinitely renewable patent-protection.

The previous result on the growth rate of the stock of products implies that:

$$g_J = \frac{\chi + \phi^{\mathrm{R}}}{1+\chi}(g_X - g_{W^{\mathrm{W}}}) + \frac{1-\phi^{\mathrm{R}}}{1+\chi}(\zeta^{\mathrm{I}} g_{A^*} + \zeta^{\mathrm{I1}} g_S) - \zeta^{\mathrm{R}} g_{A^*}.$$

The existence of a solution for our model, at all time periods, requires the number of firms in a protected industry to be bounded below asymptotically. From the previous expression for g_J, it is sufficient that $\zeta^{\mathrm{R}} g_{A^*} - \frac{1-\phi^{\mathrm{R}}}{1+\chi}\zeta^{\mathrm{I1}} g_S \leq \frac{\chi+\phi^{\mathrm{R}}}{1+\chi} g_N + \frac{1-\phi^{\mathrm{R}}}{1+\chi}\zeta^{\mathrm{I}} g_{A^*}$ for this to hold. As long as this inequality is satisfied, it may also be shown that there always exists $\underline{\underline{\mathcal{L}}}^{\mathrm{I}} > 0$ such that for $\underline{\mathcal{L}}^{\mathrm{I}} \geq \underline{\underline{\mathcal{L}}}^{\mathrm{I}}$, with probability 1, $J_t(i) \geq 2$ for all t and all $i \in \mathbb{I}_t$.

Assuming that the number of firms in a protected industry is indeed bounded below asymptotically, we show in the online appendix, section 7.4, that providing the growth rate of the productivity of newly invented products is sufficiently close to the frontier growth rate (i.e. S_t does not decline too quickly[29]), asymptotically catch-up to the frontier is instantaneous in protected industries, and the frontier growth rate is stationary. This instantaneous catch-up to the frontier means that, had we allowed for industry-specific shocks, all other protected industries would "inherit" the best industry shock, the period after it arrived. This justifies our focus on aggregate "idea" shocks. Additionally, instantaneous catch-up to the frontier means that at least asymptotically, long-run growth may be sustained even in the absence of patent-protection (i.e. when $q = 1$), as the one period in which the inventor has a first mover advantage is sufficient for their industry to surpass the existing frontier.

We summarise all the conditions previously assumed or derived in the proposition below. Since g_{A^*} is endogenous, we now replace g_{A^*} in the previous inequalities with its bounds in terms of structural parameters. In particular, we use the fact that from (1.2), (1.6), the results of the online appendix, section 7.4, and the inequalities on $d_t(i)$ derived in the online appendix, section 7.1, as long as $J_t(i) \geq 2$ asymptotically:

$$\frac{p\eta\lambda}{1-\gamma p\eta\lambda} < -\frac{1}{\gamma}\log(1-\gamma p\eta\lambda) < g_{A^*} \leq -\frac{1}{\gamma}\log\left(1-\frac{3}{2}\gamma p\lambda\right) < \frac{3}{2}p\lambda.$$

3.1.1. Proposition 1

Suppose that: $\underline{\mathcal{L}}^{\mathrm{I}} \geq \underline{\underline{\mathcal{L}}}^{\mathrm{I}}$, $\gamma > \zeta^{\mathrm{R1}} > \zeta^{\mathrm{A1}}$, $\zeta^{\mathrm{R2}} \leq \zeta^{\mathrm{A2}}$, $\phi^{\mathrm{R}} \leq \phi^{\mathrm{A}} < 1$, $\zeta^{\mathrm{R}} = \zeta^{\mathrm{R1}} - \zeta^{\mathrm{R2}} > 0$, $\zeta^{\mathrm{A}} = \zeta^{\mathrm{A1}} - \zeta^{\mathrm{A2}} > 0$, $\zeta^{\mathrm{I}} = \zeta^{\mathrm{I1}} - \zeta^{\mathrm{I2}} > 0$, $(1-\phi^{\mathrm{R}})\zeta^{\mathrm{I}} \geq (1+\chi)\zeta^{\mathrm{R}}$, $\frac{1}{\eta\lambda} < \min\{\gamma, \tau\}$, $\gamma p\lambda < \frac{2}{3}$,

$$0 \leq -g_S \leq \min\left\{\frac{\zeta^{\mathrm{R1}}-\zeta^{\mathrm{A1}}}{2\zeta^{\mathrm{R1}}-\zeta^{\mathrm{A1}}}\frac{p\eta\lambda}{1-\gamma p\eta\lambda}, \frac{(\chi+\phi^{\mathrm{R}})g_N + \left((1-\phi^{\mathrm{R}})\zeta^{\mathrm{I}} - (1+\chi)\zeta^{\mathrm{R}}\right)\frac{p\eta\lambda}{1-\gamma p\eta\lambda}}{(1-\phi^{\mathrm{R}})\zeta^{\mathrm{I1}}}\right\},$$

$$\phi^{\mathrm{R}} \leq \frac{(1+\chi)\zeta^{\mathrm{R}}\frac{p\eta\lambda}{1-\gamma p\eta\lambda}}{\max\left\{0^+, g_N - \zeta^{\mathrm{I1}} g_S - \zeta^{\mathrm{I}}\frac{p\eta\lambda}{1-\gamma p\eta\lambda}\right\}}, \qquad \phi^{\mathrm{A}} \leq \frac{(1+\chi)\zeta^{\mathrm{A}}\frac{p\eta\lambda}{1-\gamma p\eta\lambda}}{\max\left\{0^+, g_N - \zeta^{\mathrm{I1}} g_S - \zeta^{\mathrm{I}}\frac{p\eta\lambda}{1-\gamma p\eta\lambda}\right\}}, \quad \text{and}$$

$$\delta_{\mathbb{I}} > 1 - \exp\left\{-\frac{1}{1+\chi}\left[\frac{3}{2}\zeta^{\mathrm{I}} p\lambda + \zeta^{\mathrm{I1}} g_S - g_N\right]\right\}.$$

Then:

[29] A sufficient condition is that $g_S > -\frac{\zeta^{\mathrm{R1}}-\zeta^{\mathrm{A1}}}{2\zeta^{\mathrm{R1}}-\zeta^{\mathrm{A1}}} g_{A^*}$. So in the limit as $\zeta^{\mathrm{A}} \to 0$ it is sufficient that S_t is declining at less than half the rate that A_t^* is growing.

a) With probability 1, for all t and all $i \in \mathbb{I}_t$, $\frac{1}{\mu_t(i)} < \min\{\gamma, \tau\}$ and $d_t(i) > \gamma p \mu_t(i)$, so the second order and uniqueness conditions are satisfied.

b) With probability 1, $A_t^{*-\zeta^{\mathrm{R}}} \mathbb{I}_t^{\phi^{\mathrm{R}}} \to 0$ and $A_t^{*-\zeta^{\mathrm{A}}} \mathbb{I}_t^{\phi^{\mathrm{A}}} \to 0$ as $t \to \infty$, so the assumptions made in 2.2.2 are satisfied, and research and invention are indeed getting harder over time.

c) With probability 1, $J_t(i) \geq 2$ for all t and all $i \in \mathbb{I}_t$, , so the number of firms in protected industries is indeed bounded below asymptotically.

d) $g_{|\mathbb{I}|} > \log(1 - \delta_{\mathbb{I}})$, so invention does not stop asymptotically, and consequently neither does research.

e) $-g_S < \frac{\zeta^{\mathrm{R1}} - \zeta^{\mathrm{A1}}}{2\zeta^{\mathrm{R1}} - \zeta^{\mathrm{A1}}} g_{A^*}$, so asymptotically catch-up to the frontier is instantaneous in protected industries.

f) Research and appropriation are not performed in non-protected industries asymptotically.

g) The asymptotic growth rate of consumption and output is equal to $g_N + g_{A^*}$, where $g_{A^*} > \frac{p\eta\lambda}{1 - \gamma p \eta \lambda} > 0$.

Furthermore, the set of parameters satisfying the conditions of this proposition is of strictly positive measure. Hence, the model generates fully endogenous growth, without knife-edge assumptions.

3.2. Short and medium run behaviour

We now turn to an examination of the qualitative behaviour of our model over the short to medium run, something we will assess via simulations.

If the number of firms in protected industries were asymptotically infinite, then these simulations would tell us nothing about the consequences of the variations in this number that we might see non-asymptotically. Therefore, it will be helpful if it is additionally the case that this number is finite and stationary asymptotically. To guarantee this will, unfortunately, require a knife-edge assumption, namely that $\frac{E_t^{\mathrm{R}}(\sup \mathbb{I}_t)}{E_t^{\mathrm{I}}}$ is stationary. However, this assumption should just be viewed as a trick to recover some of the model's non-asymptotic dynamics, even asymptotically. Unlike with knife-edge growth models whereby relatively slight departures from the special parameter values result in growth that could not possibly explain our observed stable exponential growth, here, away from the knife-edge case we will have slowly decreasing mark-ups, consistent with Ellis's (2006) evidence of a persistent decline in UK whole economy mark-ups over the last thirty years and Kim's (2010) evidence of non-stationarity in mark-ups.

We assume then that the difficulty of producing a prototype (S_t) is such that $\frac{E_t^{\mathrm{R}}(\sup \mathbb{I}_t)}{E_t^{\mathrm{I}}}$ is stationary.[30] In fact, without loss of generality we may assume $\frac{E_t^{\mathrm{R}}(\sup \mathbb{I}_t)}{E_t^{\mathrm{I}}} \equiv 1$, since the only time $\frac{E_t^{\mathrm{R}}(\sup \mathbb{I}_t)}{E_t^{\mathrm{I}}}$ enters the model's equations is when multiplied by $\mathcal{L}_t^{\mathrm{I}}$, which is already an arbitrary stationary stochastic process. Under this assumption, as $t \to \infty$ the behaviour of our model tends towards stationarity in the key variables. We simulate this asymptotically stationary model. Asymptotically non-protected industries will perform no research or appropriation, so their entry cost to post-entry industry profits ratio is tending to zero, meaning their number of firms will tend to infinity as $t \to \infty$. This is in line with our motivating intuition that excess entry in non-protected industries kills research and appropriation incentives. The full set of equations of the de-trended model is given in the online appendix, section 7.5. The definition of equilibrium here is entirely standard.

[30] Recall that we require $g_S > -\frac{\zeta^{\mathrm{R1}} - \zeta^{\mathrm{A1}}}{2\zeta^{\mathrm{R1}} - \zeta^{\mathrm{A1}}} g_{A^*}$ for long run growth. Hence, under the assumption that $A_t^{*-\zeta^{\mathrm{R}}} I_t^{\phi^{\mathrm{R}}} S_t^{\zeta^{\mathrm{I1}}} A_t^{*\zeta^{\mathrm{I}}} I_{t-1}^{\chi}$ is stationary we require: $\frac{1}{\zeta^{\mathrm{I1}}}\left[\left(\zeta^{\mathrm{I}} - \frac{1+\chi}{1-\phi^{\mathrm{R}}}\zeta^{\mathrm{R}}\right) g_{A^*} + \frac{\chi + \phi^{\mathrm{R}}}{1-\phi^{\mathrm{R}}} g_N\right] < \frac{\zeta^{\mathrm{R1}} - \zeta^{\mathrm{A1}}}{2\zeta^{\mathrm{R1}} - \zeta^{\mathrm{A1}}} g_{A^*}$, which always holds for sufficiently large ζ^{I1}.

When $\lambda = \nu = \gamma = 1$, it may be shown analytically that the equations determining the model's steady-state have at most two solutions with more than one firm in each industry. However, only one of these two solutions exists for large values of $\mathcal{L}^{\mathrm{I}}$, i.e. when invention is costly. Since away from our knife edge assumption, invention may be getting harder over time faster than research (due to congestion effects say), any solution that only exists for small values of $\mathcal{L}^{\mathrm{I}}$ is non-feasible. Our numerical investigations suggest that the model always has at most these two equilibria, and that always at most one of them exists for large values of $\mathcal{L}^{\mathrm{I}}$.[31] However, at the chosen parameters, the model has a unique solution, which will exist for arbitrarily high values of $\mathcal{L}^{\mathrm{I}}$.

We fix all of the model's other parameters, except $\mathcal{L}^{\mathrm{I}}$, to the values we estimated in a medium-scale version of the model in TODO. $\mathcal{L}^{\mathrm{I}}$ is set such that the number of firms in patent-protected industries in this model is equal to that in the estimated extended model. The full parameterisation is reported in the online appendix TODO. We note that β is set to 0.96 consistent with an interpretation as an annual model, given our focus on medium and lower frequency phenomena.

In Figure 3 TODO: OUT OF DATE we present the nonlinear perfect foresight impulse responses that result from IID (hence non-persistent) shocks to population growth ($G_{N,t}$) and "ideas" (Z_t), in the fully nonlinear model. [32] We set the magnitude of the idea shock to 1%, and choose the magnitude of the population shock to give a similar productivity response after 5 years.[33] Each graph is given in terms of per cent deviations from the value the variable would have taken had the shock never arrived, and the horizontal axis shows time in years, though this remains a quarterly model.

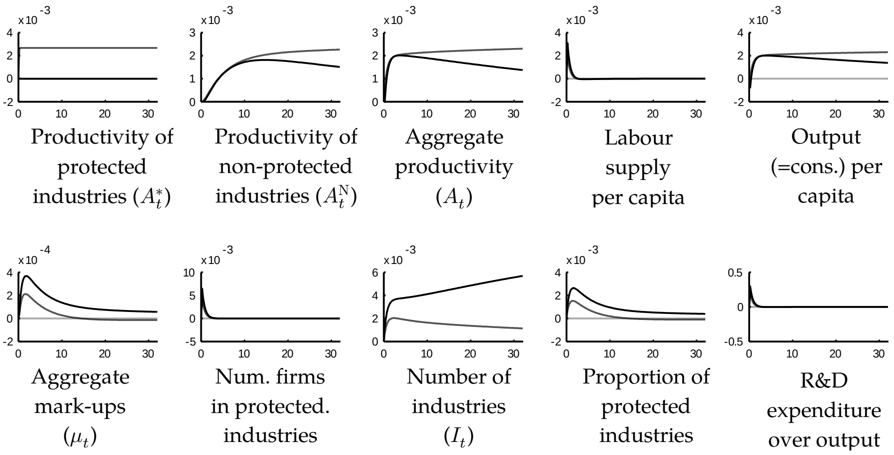

Figure 3: TODO OUT OF DATE Impulse responses from population (solid) and idea (dashed) shocks.
(Vertical axes are in percent, horizontal axes are in years.)

The principle mechanism of our paper is illustrated most clearly by the population growth rate shock, shown by a solid line in each graph. (We do not wish to advance population shocks as a key driver of business cycles though, since real rigidities will significantly reduce their impact.) Following a permanent increase in population, demand is permanently higher, so, in the long run, the number of industries must grow to balance this out. Given sufficiently inelastic labour supply, this long run increase in the measure of industries requires a short-run substitution of labour from production to invention, pushing down consumption and pushing up wages, and so moderating the rate at which invention will grow. Consequently, in the short run

[31] It may be shown analytically that the steady state of the complete model may always be solved for by solving a single nonlinear equation, which was always concave for all the parameters we examined.
[32] This was performed using Dynare (Adjemian et al. 2011).
[33] This required a 0.01% shock to $G_{N,t}$.

some of the additional demand is absorbed by fluctuations in the number of firms in each industry. Without this additional margin of adjustment, this shock would have led to a large increase in average firm sizes, with a consequent increase in the frontier growth rate and counter-factually large unit root in output.

Despite the tiny movement in frontier productivity (less than 0.000001%), there is still however a substantial movement in aggregate productivity in the medium-term. Following the shock, more new products are being invented each period, meaning that a greater proportion of industries are relatively new, and so a greater proportion are patent-protected. But because patent-protected industries have such strong incentives to catch-up to the frontier, patent-protected industries are more productive than non-protected ones, so an increase in the proportion of industries that are patent protected means an increase in aggregate productivity. Patent-protected industries also have higher mark-ups due to the cost of paying licence fees, enabling our model to generate pro-cyclical mark-ups.[34] Indeed, since there is so much persistence in productivity coming from patent protection, as in the data, movements in mark-ups lead movements in output, bringing us close to the observed cross correlation even in this toy model.

Fluctuating invention rates also drive the model's response to any other shock that might be considered, not least the idea shock, shown by the dashed line in each graph. Initially, an idea shock just increases the productivity of patent-protected industries. This also makes them relatively more profitable, enabling patent holders to extract higher rents, and leading to an increase in invention with a corresponding further increase in aggregate productivity. Over time, patent protected industries fall out of patent-protection, carrying their higher productivity with them, and thus increasing the average productivity of non-protected firms too. Consequently, aggregate productivity slowly rises towards its permanently higher long run level. However, even with this reasonably large research productivity shock (1%), frontier productivity still only rises by less than TODO 0.005%, consistent with very low variance at frequency zero.

4. Medium-frequency cycles in a modern real business cycle framework

TODO a lot in the below…

We go on to embed our widget-producing sector in a richer annual real business cycle model with capital and consumer durables, the latter modelled as an input to home production. In order to capture the effects of the long-run risk induced by medium frequency cycles, following Kung and Schmid (forthcoming), we include Epstein-Zin (1989) preferences. We also allow for skill accumulation, which, in the model, provides access to higher tiers of a segmented labour market. The final good will now be produced by combining labour, capital and widgets, where the labour required to produce the final good is of a less skilled variety than that required to produce widgets or perform research, appropriation or invention. This segmentation in labour markets will partially insulate the widget sector from business cycles, dampening an otherwise overly powerful amplification mechanism.

[34] Pavlov and Weder (2012) also develop a business cycle model capable of generating pro-cyclical mark-ups, via the changing importance of different types of *buyers* over the business cycle. The properties of these buyers are exogenous in their model however, whereas the properties of the different types of *sellers* that drive our results are endogenous.

There will be several different production functions in the model, but all of them will take the same form. We introduce this form and its properties now to save repeated exposition. In particular, for any string of letters S, let $f_S(\cdot,\cdot,\cdot,\cdot)$ be the production function given by:

$$f_S(x_{1,t}, x_{2,t}, x_{1,t-1}, x_{2,t-1}) = \frac{\Gamma_t^S x_{1,t}^{\alpha_S} x_{2,t}^{1-\alpha_S}}{\sqrt{1 + \varsigma_S \left[\log\left(\frac{x_{1,t}x_{2,t-1}}{x_{2,t}x_{1,t-1}}\right)\right]^2}},$$

where Γ_t^S is a stationary stochastic process representing fluctuations in technology, $\alpha_S \in (0,1)$ and $\varsigma_S \geq 0$. This combines a standard Cobb-Douglas production function with an adjustment cost (controlled by ς_S) that penalises changes in the $\frac{x_{1,t}}{x_{2,t}}$ ratio. The specific form of the adjustment cost was chosen to ensure that it was differentiable and symmetric in $x_{1,t}$ and $x_{2,t}$, and to ensure that f_S was homogeneous of degree one, positive and strictly increasing in $x_{1,t}$ and $x_{2,t}$ (for $\varsigma_S < 4\min\{\alpha_S^2, (1-\alpha_S)^2\}$) and concave in $x_{1,t}$ and $x_{2,t}$ (for $\varsigma_S \leq 3\min\{\alpha_S^2, (1-\alpha_S)^2\}$). The adjustment cost ensures that although the elasticity of substitution between $x_{1,t}$ and $x_{2,t}$ is 1 in the long run, in the short run it is less than one.[35] This specification enables us to match the evidence on short-run complementarity, while remaining consistent with long-run growth, even when the prices of both factors are non-stationary. It may also be viewed as a reduced form approximation to the production technique choice model of Leon-Ledesma and Satchi (2011).

4.1. Firms

There are three sectors of price-taking firms in the model, namely those producing capital, durable and market consumption goods. The representative firm producing good G ∈ {K,D,CM} chooses their period t labour and widget demands, L_t^G and X_t^G respectively, and their period t capital investment, $I_{K,t}^G$, to maximise the value function:

$$V_t^G = P_t^G Y_t^G - W_t L_t^G - X_t^G - P_t^K I_{K,t}^G + \beta \mathbb{E}_t \Xi_{t+1} V_{t+1}^G,$$

where P_t^G is the market price of good G, Y_t^G is the output of good G, P_t^K is the market price of capital, $Y_t^G = f_O\left(\mathfrak{K}_t^G, \mathfrak{L}_t^G, \mathfrak{K}_{t-1}^G, \mathfrak{L}_{t-1}^G\right)$, $\mathbb{E}\Gamma_t^O = 1$ (for identification), $\mathfrak{L}_t^G = f_{GL}\left(X_t^G, L_t^G, X_{t-1}^G, L_{t-1}^G\right)$ and:

$$\mathfrak{K}_t^G = \left((1-\theta_K)K_t^G + \theta_K K_{t-1}^G\right)\left(1 - \left(\frac{I_{K,t}^G}{K_t^G}\right)^{\frac{1}{v_{K,1}}}\right)^{v_{K,2}},$$

where K_t^G is the stock of sector specific capital, $v_{D,1}, v_{D,2} \in [0,1]$ determine the strength of the capital good adjustment costs, and θ_D controls the strength of the "time-to-build" friction, as one period time to build may be excessive with annual data. The intuition for this form of adjustment costs stems from the fact that the installation of new capital gets in the way of production, reducing the productivity of the existing stock. Imposing capital good adjustment costs in this way means the capital accumulation equation remains standard, enabling us to use NIPA measures of the capital stock. In particular, K_t^G evolves according to:

$$K_t^G = (1-\delta_{K,t})K_{t-1}^G + I_{K,t}^G,$$

where $\delta_{K,t}$ is a stationary stochastic process giving the rate of capital depreciation. Note that all three categories of goods are produced using the same outer production function. This facilitates parameter identification, and allows for a common productivity shock across all three categories. However, the differing inner production functions mean that the various goods may have different growth rates, facilitating the matching of their differing price trends.

[35] In particular, when $\frac{x_{1,t}x_{2,t-1}}{x_{2,t}x_{1,t-1}} = 1$, the elasticity of substitution is $1 - \frac{\varsigma_V}{\alpha_V(1-\alpha_V)+\varsigma_V}$.

4.2. Families and households

We assume there is a unit mass of families, each of which contains N_t households in period t.[36] At any point in time, each household will be one of three types: untrained, trained or highly trained. We assume that untrained workers cannot work in either production of final goods (i.e. the market consumption good, the durable good or the investment good), or in the production of widgets, whereas trained workers may work in the production of final goods (but not widgets), and highly trained workers may work in production of either final goods, or widgets. In period t, the family offers households a different bundle of market consumption, durable good loan, widgets and hours depending on their reported type. The offered bundles are shown in Table 1.

Reported type	Untrained	Trained	Highly trained
Given units of the market consumption good	C_t^{MU}	C_t^{MT}	C_t^{MH}
Lent units of durable goods	D_t^{U}	D_t^{T}	D_t^{H}
Given units of widgets	X_t^{U}	X_t^{T}	X_t^{H}
Required hours in the production of final goods	0	L_t^{ST}	0
Required hours in the production of widgets	0	0	L_t^{SH}

Table 1: Offered bundles

We presume that individual households have no direct access to labour or goods markets, and so they cannot bypass the family and arrive at some other consumption bundle. Furthermore, in equilibrium, the bundles offered will be incentive compatible, so trained workers will have no incentive to pretend to be untrained, and highly trained will have no incentive to pretend to be non-highly trained or untrained.

We assume that households do not value leisure per se. Instead, following Campbell and Ludvigson (2001), we suppose that whenever they are not supplying labour to the market, they are instead engaged in home production, which produces an output they value from labour, durables and widgets, and which only they may consume.[37] All households have an endowment of $\overline{L}_t$ total usable hours in period t, where $\overline{L}_t$ is a stationary stochastic process capturing fluctuations in the amount of time that may be used in home or market production. This shock process captures various technological factors, such as those determining how much time is spent travelling to work, or the practicality of working after dark. Household i's final consumption, $C_t^{\mathrm{F}}(i) = f_{\mathrm{HF}}\big(C_t^{\mathrm{M}}(i), C_t^{\mathrm{H}}(i), C_{t-1}^{\mathrm{M}}, C_{t-1}^{\mathrm{H}}\big)$ where $C_t^{\mathrm{M}}(i)$ and $C_t^{\mathrm{H}}(i)$ are their period t consumption of the market and home goods, respectively, and where C_{t-1}^{M} and C_{t-1}^{H} are the corresponding averages across all households in the family in period $t-1$. The use of family averages here is done chiefly for technical reasons, to ensure there are no household specific state variables; however, it may also be justified via "keeping up with the Joneses" style arguments.

Each household has access to the same two-stage technology for producing the home good. They first produce a home labour bundle, $\mathfrak{L}_t^{\mathrm{H}}(i) = f_{\mathrm{HL}}\big(X_t(i), \overline{L}_t - L_t(i), X_{t-1}, \overline{L}_t - L_{t-1}\big)$ where $X_t(i)$ is the number of widgets given to them by the family, $L_t(i) \le \overline{L}_t$ is the amount of labour they are required to supply to the market, and X_{t-1} and L_{t-1} are the corresponding family averages in the previous period. They then combine this labour bundle with the durable goods lent to them by the family, $D_t(i)$, to make the home good, $C_t^{\mathrm{H}}(i) =$

[36] This terminology is borrowed from Christiano, Trabandt, and Walentin (2010) which also inspired the structure.

[37] We were inspired to use the Campbell and Ludvigson (2001) trick in an Epstein-Zin context by Rudebusch and Swanson (2012).

$f_{\mathrm{HC}}\big(D_t(i), \mathfrak{L}_t^{\mathrm{H}}(i), D_{t-1}, \mathfrak{L}_{t-1}^{\mathrm{H}}\big)$, where $\mathbb{E}\Gamma_t^{\mathrm{H}} = 1$ (for identification). Our motivation for avoiding adjustment costs in the household production functions should be readily apparent at this point, since otherwise we would have had per-household state variables. As it is however, all households of the same type will consume the same quantity. Henceforth then, we drop the i's indexing households, and label them by type, so, for example, C_t^{FU}, C_t^{FT} and C_t^{FH} are the final consumption of untrained, trained and highly trained households, respectively.

In equilibrium, the market wage in the widget production sector, W_t^{W}, will be higher than the market wages in the production of other goods, W_t^{O}. As a result, it will be most costly for the family to provide utility to the highly trained households. It will also be more expensive to provide utility to trained households than untrained ones, because the market wage of the latter is effectively zero. We assume the family cares about the sum of the utility of its constituent households, and so the incentive compatibility constraints will bind. This further simplifies the family's problem. In particular, if there are U_t untrained households, T_t trained ones and H_t highly trained ones in period t, where $N_t = U_t + T_t + H_t$, with Epstein-Zin preferences on the behalf of households,[38] the family will just maximise:

$$V_t^{\mathrm{F}} = N_t \frac{\varsigma_U}{\varsigma_U - 1} C_t^{\mathrm{F}\,\frac{\varsigma_U - 1}{\varsigma_U}} + \beta \left(\mathbb{E}_t V_{t+1}^{\mathrm{F}\;(1-\varrho)\frac{\varsigma_U}{\varsigma_U - 1}} \right)^{\frac{1}{1-\varrho}\frac{\varsigma_U - 1}{\varsigma_U}},$$

where ϱ is the coefficient of relative risk aversion and ς_U is the elasticity of intertemporal substitution, subject to the constraint that $C_t^{\mathrm{F}} = C_t^{\mathrm{FU}} = C_t^{\mathrm{FT}} = C_t^{\mathrm{FH}}$, the family budget constraint, and the laws of motion for the family state variables.

The family has several investment opportunities available to them. Firstly, they may purchase one period, zero net supply, widget denominated bonds, paying an interest rate R_t. Secondly, they may buy dollar denominated bonds which pay an interest rate $R_t^{\$}$, for a price of $P_t^{\$}$ widgets. Although this is a model without nominal rigidities, changes in the inflation risk premium will still have real effects. For simplicity, we set $P_t^{\$} := \frac{P_t^{\mathrm{CM}}}{P_t^{\mathrm{CM}/\$}}$ where P_t^{CM} is the price of the market consumption good in units of widgets, and $P_t^{\mathrm{CM}/\$}$ is a reduced form statistical model of the dollar price of market consumption, possibly correlated with other observable variables.

Thirdly, the family may invest in durable goods. Their per-capita stock, D_t, evolves according to:

$$N_t D_t = (1 - \delta_{D,t}) N_{t-1} D_{t-1} + I_{D,t},$$

where $\delta_{D,t}$ is a stationary stochastic process giving the rate of durable good depreciation, and $I_{D,t}$ is their level of investment in durables. Much as with firm's capital, the stock of durables available for use by the family in period t is given by $\big((1-\theta_D)D_t + \theta_D D_{t-1}\big)\left(1 - \left(\frac{I_{D,t}}{D_t}\right)^{\frac{1}{v_{D,1}}}\right)^{v_{D,2}}$, where $v_{D,1}, v_{D,2} \in [0,1]$ determine the strength of these durable good adjustment costs, and θ_D controls the strength of the "time-to-build" friction.

Fourthly, the family may invest in the skills of its constituent households. We assume that training a household requires labour supply from both the trainer and the trainee. We also assume the existence of congestion externalities that make it difficult to train many people in one period. Suppose the family decides to train $I_{T,t}$ households from untrained to trained in period t, and $I_{H,t}$ households from trained to highly-

[38] To derive the given expression through aggregation we need to assume that households care about the sum of their own future utility and that of the new households to which they give birth.

trained in that period. For simplicity, we posit that those households change type immediately at the start of the period. However, a total of $\frac{\xi_T I_{T,t}}{\left(1-\left(\frac{I_{T,t}}{T_t}\right)^{\frac{1}{v_{T,1}}}\right)^{v_{T,2}}}$ extra units of trained labour and $\frac{\xi_H I_{H,t}}{\left(1-\left(\frac{I_{H,t}}{H_t}\right)^{\frac{1}{v_{H,1}}}\right)^{v_{H,2}}}$ extra units of highly trained labour are required during the period, representing both the trainers' and trainees' training effort. The stocks of trained and highly trained labour evolve according to:

$$T_t = (1-\delta_{T,t})T_{t-1} + I_{T,t},$$
$$H_t = (1-\delta_{H,t})H_{t-1} + I_{H,t},$$

where $\delta_{T,t}$ and $\delta_{H,t}$ are stationary stochastic process determining the rate of skill depreciation.

Finally, the family may invest in shares in several classes of firms, namely: the firms within each industry in the widget producing sector, the firms holding patents in that sector, the firms producing capital goods, the firms producing durable goods, and the firms producing market consumption goods. Combining these various investment opportunities leads to the family's period t budget constraint:

$$W_t^{\mathrm{O}}\left[T_t L_t^{\mathrm{T}} - \frac{\xi_T I_{T,t}}{\left(1-\left(\frac{I_{T,t}}{T_t}\right)^{\frac{1}{v_{T,1}}}\right)^{v_{T,2}}}\right] + W_t^{\mathrm{W}}\left[H_t L_t^{\mathrm{H}} - \frac{\xi_H I_{H,t}}{\left(1-\left(\frac{I_{H,t}}{H_t}\right)^{\frac{1}{v_{H,1}}}\right)^{v_{H,2}}}\right] + R_{t-1}B_{t-1} + \frac{R_{t-1}^{\$}B_{t-1}^{\$}}{P_t^{\$}} + \Pi_t$$
$$= B_t + P_t^{\$}B_t^{\$} + U_t X_t^{\mathrm{U}} + T_t X_t^{\mathrm{T}} + H_t X_t^{\mathrm{H}} + P_t^{\mathrm{CM}}\left[U_t C_t^{\mathrm{MU}} + T_t C_t^{\mathrm{MT}} + H_t C_t^{\mathrm{MH}}\right] + P_t^{\mathrm{D}} I_{D,t},$$

where B_t is their widget denominated bond holdings, $B_t^{\$}$ is their dollar denominated bond holdings and Π_t is their dividend income.

Consequently, the family's stochastic discount factor is given by: $\beta\Xi_{t+1} = \beta\mathcal{D}_{t+1}\frac{\alpha_{t+1}}{\alpha_t}$, where $\mathcal{D}_{t+1} := \left[\frac{V_{t+1}^{\mathrm{F}}}{\left(\mathbb{E}_t V_{t+1}^{\mathrm{F}\,(1-\varrho)\frac{\varsigma_U}{\varsigma_U-1}}\right)^{\frac{1}{1-\varrho}\frac{\varsigma_U-1}{\varsigma_U}}}\right]^{(1-\varrho)\frac{\varsigma_U}{\varsigma_U-1}-1}$ and where α_t is the Lagrange multiplier on the family's budget constraint, which is given by:

TODO UPDATE!

$$\alpha_t := \frac{\alpha_{\mathrm{HF}}}{P_t^{\mathrm{CM}}}\left(\frac{N_t C_t^{\mathrm{F}}}{U_t C_t^{\mathrm{MU}} + T_t C_t^{\mathrm{MT}} + H_t C_t^{\mathrm{MH}}}\right)\left[\left(C_t^{\mathrm{F}} - hC_{t-1}^{\mathrm{F}}\right)^{-\frac{1}{\varsigma_U}} - \beta h\mathbb{E}_t\mathcal{D}_{t+1}G_{N,t+1}\left(C_{t+1}^{\mathrm{F}} - hC_t^{\mathrm{F}}\right)^{-\frac{1}{\varsigma_U}}\right].$$

The other first order conditions of the family are given in appendix TODO.

4.3. Market clearing

In equilibrium, all markets clear, and household, implying:

$$X_t = U_t X_t^{\mathrm{U}} + T_t X_t^{\mathrm{T}} + H_t X_t^{\mathrm{H}} + X_t^{\mathrm{K}} + X_t^{\mathrm{D}} + X_t^{\mathrm{CM}},$$
$$T_t L_t^{\mathrm{T}} = \frac{\xi_T I_{T,t}}{\left(1-\left(\frac{I_{T,t}}{T_t}\right)^{\frac{1}{v_{T,1}}}\right)^{v_{T,2}}} + L_t^{\mathrm{K}} + L_t^{\mathrm{D}} + L_t^{\mathrm{CM}},$$
$$H_t L_t^{\mathrm{H}} = \frac{\xi_H I_{H,t}}{\left(1-\left(\frac{I_{H,t}}{H_t}\right)^{\frac{1}{v_{H,1}}}\right)^{v_{H,2}}} + \frac{\mathcal{L}_t^{\mathrm{I}}}{E_t^{\mathrm{I}}}\left[|\mathbb{I}_t| - (1-\delta_{\mathbb{I},t})|\mathbb{I}_{t-1}|\right] + \int_{i\in\mathbb{I}_t}\left(L_t^{\mathrm{R}}(i) + L_t^{\mathrm{A}}(i) + L^{\mathrm{F}}\right)J_t(i)\,di$$
$$+ W_t^{\mathrm{W}^{-\frac{1+\lambda}{\lambda}}}\frac{X_t}{|\mathbb{I}_{t-1}|}\int_{i\in\mathbb{I}_{t-1}}\frac{1}{A_t(i)}\left(\frac{A_t(i)}{1+\mu_{t-1}(i)}\right)^{\frac{1+\lambda}{\lambda}}di,$$
$$Y_t^{\mathrm{K}} = I_{K,t}^{\mathrm{K}} + I_{K,t}^{\mathrm{D}} + I_{K,t}^{\mathrm{CM}},$$

$$Y_t^{\mathrm{D}} = I_{D,t},$$
$$Y_t^{\mathrm{CM}} = U_t C_t^{\mathrm{MU}} + T_t C_t^{\mathrm{MT}} + H_t C_t^{\mathrm{MH}}, \quad \text{and}$$
$$N_t D_t = U_t D_t^{\mathrm{U}} + T_t D_t^{\mathrm{T}} + H_t D_t^{\mathrm{H}}.$$

4.4. Structural validation

TODO

1. Population
2. Total capital
3. Total durables
4. R&D stock
5. Price of consumption
6. Price of capital investment
7. Price of durable investment
8. Price of R&D investment
9. Capital investment
10. Durable investment
11. R&D investment
12. Stock market value
13. Nominal interest rates
14. Total consumption
15. Total labour supply
16. Total labour compensation
17. Labour force participation

1. Z_t Stock market value
2. $\tilde{\delta}_{\mathbb{I},t}$ R&D stock
3. $\mathcal{L}_t^{\mathrm{I}}$ Price of R&D investment
4. N_t Population
5. $\delta_{K,t}$ Total capital
6. $\delta_{D,t}$ Total durables
7. $\delta_{T,t}$ Labour force participation
8. $\delta_{H,t}$ R&D investment
9. Γ_t^{O} Capital investment
10. Γ_t^{KL} Price of capital investment
11. Γ_t^{DL} Price of durable investment
12. Γ_t^{CML} Durable investment
13. $\overline{L}_t$ Total labour supply
14. Γ_t^{HF} Nominal interest rates
15. Γ_t^{HL} Total labour compensation
16. Γ_t^{HC} Total consumption
17. $P_t^{\mathrm{CM/\$}}$ Price of consumption

4.5. Reduced form validation

TODO

5. Conclusion

TODO UPDATE

Many have expressed the worry that "the apparent fit of the DSGE model [has] more to do with the inclusion of suitable exogenous driving processes than with the realism of the model structure itself"[39]. In this paper, we have demonstrated that if productivity is endogenized through research, appropriation and invention then even a frictionless RBC model is capable of generating rich persistent dynamics from uncorrelated shocks, thanks to fluctuations in the proportion of industries that are producing patent-protected products. Furthermore, this improvement in the model's propagation mechanism does not come at the expense of implausibly large trend breaks in output following shocks, counter-factual movements in mark-ups, or the use of a growth model that we can reject thanks to the absence of strong scale effects in the data. In all of these respects, then, our model presents a substantial advance on the prior literature. We went on to embed our core model within a modern real business cycle framework, showing that this enables just a few shocks to explain much of the data at both business and medium frequencies.

Our model suggests that a switch to indefinite patent protection would result in significant welfare improvements. Such a switch would both permanently increase the level of aggregate productivity, and substantially lessen its variance and persistence, while only slightly increasing mark-ups and efficiency losses due to research duplication. Indeed, it may be shown that in our model increasing patent protection even slightly increases growth rates, as industry profits are decreasing in aggregate productivity, and so with indefinite patent protection each (protected) industry has fewer firms meaning higher mark-ups and higher research. For similar reasons, increasing product differentiation through (for example) trademark law would increase growth rates in our model, as indeed would increasing patent-holder bargaining power.

However, it is clear that the structure of our model has "stacked-the-deck" in favour of finding a beneficial role for patent protection and monopolistic power. Patents in our model are less broad than in the real world, and they do not hinder future research or invention. One minimal conclusion we can draw on patent protection is that product patents should at least be long enough that by the end of patent protection, production process have reached frontier productivity. In our model, this time goes to zero asymptotically. A less radical policy change might be to grant temporary extensions to patents that would otherwise expire during a recession. We intend to explore the full policy implications of this model in future work.

6. References

Adjemian, Stéphane, Houtan Bastani, Michel Juillard, Ferhat Mihoubi, George Perendia, Marco Ratto, and Sébastien Villemot. 2011. *Dynare: Reference Manual, Version 4*. Dynare Working Papers. CEPREMAP.

[39] Del Negro et al. (2007) paraphrasing Kilian (2007).

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics* 120 (2) (May): 701–728.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60 (2) (March): 323–351.

Autor, David H, Lawrence F Katz, and Melissa S Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics* 90 (2): 300–323.

Bachmann, Ruediger, and Christian Bayer. 2009. *The Cross-Section of Firms over the Business Cycle: New Facts and a DSGE Exploration*. CESifo Group Munich.

Balasubramanian, Natarajan, and Jagadeesh Sivadasan. 2011. "What Happens When Firms Patent? New Evidence from U.S. Economic Census Data." *Review of Economics and Statistics* 93 (1) (July 7): 126–146.

Bental, Benjamin, and Dan Peled. 1996. "The Accumulation of Wealth and the Cyclical Generation of New Technologies: A Search Theoretic Approach." *International Economic Review* 37 (3) (August): 687–718.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz. 2011. *Endogenous Entry, Product Variety, and Business Cycles*. CEPR Discussion Papers. C.E.P.R. Discussion Papers.

———. 2012. "Endogenous Entry, Product Variety, and Business Cycles." *Journal of Political Economy* 120 (2) (April 1): 304–345.

Bils, Mark. 1998. "Discussion of Basu (1998)." *Conference Series ; [Proceedings]* (Jun).

Bonanno, Giacomo. 1988. "Oligopoly Equilibria When Firms Have Local Knowledge of Demand." *International Economic Review* 29 (1) (February 1): 45–55.

Boulhol, Hervé. 2007. "The Convergence of Price–cost Margins." *Open Economies Review* 19 (2) (May): 221–240.

Broda, Christian, and David E Weinstein. 2010. "Product Creation and Destruction: Evidence and Price Implications." *American Economic Review* 100 (3) (June): 691–723.

Campbell, Jeffrey R. 1998. "Entry, Exit, Embodied Technology, and Business Cycles." *Review of Economic Dynamics* 1 (2) (April): 371–408.

Campbell, John Y, and Sydney Ludvigson. 2001. "Elasticities of Substitution in Real Business Cycle Models with Home Protection." *Journal of Money, Credit and Banking* 33 (4) (November): 847–75.

Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright. 2006. *Historical Statistics of the United States Millennial Edition Online*. Table Ca9-19: Gross domestic product: 1790-2002 [Continuous annual series]. New York: Cambridge University Press.

Christiano, Lawrence J., Mathias Trabandt, and Karl Walentin. 2010. *Involuntary Unemployment and the Business Cycle*. National Bureau of Economic Research, Inc.

Cochrane, John H. 1988. "How Big Is the Random Walk in GNP?" *Journal of Political Economy* 96 (5) (October 1): 893–920.

Comin, Diego. 2009. "On the Integration of Growth and Business Cycles." *Empirica* 36 (2) (May 1): 165–176.

Comin, Diego, and Mark Gertler. 2006. "Medium-Term Business Cycles." *American Economic Review* 96 (3) (June): 523–551.

Comin, Diego, Mark Gertler, and Ana Maria Santacreu. 2009. *Technology Innovation and Diffusion as Sources of Output and Asset Price Fluctuations*. National Bureau of Economic Research, Inc.

Comin, Diego, and Sunil Mulani. 2009. "A Theory of Growth and Volatility at the Aggregate and Firm Level." *Journal of Monetary Economics* 56 (8) (November): 1023–1042.

Dixit, Avinash K., and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *The American Economic Review* 67 (3) (June): 297–308.

Duarte, Margarida, and Diego Restuccia. 2009. "The Role of the Structural Transformation in Aggregate Productivity". Working Papers tecipa-348. University of Toronto, Department of Economics.

Eisfeldt, Andrea L., and Adriano A. Rampini. 2006. "Capital Reallocation and Liquidity." *Journal of Monetary Economics* 53 (3) (April): 369–399.

Ellis, Colin. 2006. *Elasticities, Markups and Technical Progress: Evidence from a State-Space Approach*. SSRN.

Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57 (4) (July 1): 937–969.

Ethier, Wilfred J. 1982. "National and International Returns to Scale in the Modern Theory of International Trade." *The American Economic Review* 72 (3) (June): 389–405.

Francois, Patrick, and Huw Lloyd-Ellis. 2008. "Implementation Cycles, Investment, and Growth." *International Economic Review* 49 (3) (August 1): 901–942.

———. 2009. "Schumpeterian Cycles with pro-Cyclical R&D." *Review of Economic Dynamics* 12 (4) (October): 567–591.

Groth, Christian, Karl-Josef Koch, and Thomas Steger. 2009. "When Economic Growth Is Less than Exponential." *Economic Theory* (July 4).

Hart, Oliver, and John Moore. 1994. "A Theory of Debt Based on the Inalienability of Human Capital." *The Quarterly Journal of Economics* 109 (4) (November 1): 841–879.

Holden, Tom. 2011. *Products, Patents and Productivity Persistence: A DSGE Model of Endogenous Growth*. Dynare Working Papers. CEPREMAP.

———. 2013. "Reconciling near Trend-Stationary Growth with Medium-Frequency Cycles."

Horii, Ryo. 2011. "Wants and Past Knowledge: Growth Cycles with Emerging Industries." *Journal of Economic Dynamics and Control* In Press, Accepted Manuscript.

Jaimovich, Nir. 2007. "Firm Dynamics and Markup Variations: Implications for Sunspot Equilibria and Endogenous Economic Fluctuations." *Journal of Economic Theory* 137 (1) (November): 300–325.

Kehrig, Matthias. 2011. *The Cyclicality of Productivity Dispersion*. Center for Economic Studies, U.S. Census Bureau.

Kilian, Lutz. 2007. "Comment." *Journal of Business & Economic Statistics* 25. Journal of Business & Economic Statistics: 156–159.

Kim, Bae-Geun. 2010. "Identifying a Permanent Markup Shock and Its Implications for Macroeconomic Dynamics." *Journal of Economic Dynamics and Control* (April 14).

Kung, Howard, and Lukas Schmid. forthcoming. "Innovation, Growth, and Asset Prices." *Journal of Finance*

Leon-Ledesma, Miguel A., and Mathan Satchi. 2011. *The Choice of CES Production Techniques and Balanced Growth*. Studies in Economics. Department of Economics, University of Kent.

Levin, Richard C., Wesley M. Cohen, and David C. Mowery. 1985. "R & D Appropriability, Opportunity, and Market Structure: New Evidence on Some Schumpeterian Hypotheses." *The American Economic Review* 75 (2) (May): 20–24.

Li, Chol-Won. 2000. "Endogenous vs. Semi-Endogenous Growth in a Two-R&D-Sector Model." *The Economic Journal* 110 (462) (March 1): C109–C122.

Mansfield, Edwin. 1993. "The Diffusion of Industrial Robots in Japan and the United States." *Research Policy* 22 (2) (April): 105.

Matsuyama, Kiminori. 1999. "Growing Through Cycles." *Econometrica* 67 (2) (March): 335–347.

Del Negro, Marco, Frank Schorfheide, Frank Smets, and Rafael Wouters. 2007. "Rejoinder." *Journal of Business and Economic Statistics* 25 (2): 159.

Nekarda, Christopher, and Valerie Ramey. 2010. "The Cyclical Behavior of the Price-Cost Markup."

Nuño, Galo. 2008. *Schumpeterian Foundations of Real Business Cycles*. Working Papers. International Economics Institute, University of Valencia.

———. 2009. *Technology, Convergence and Business Cycles*. Banco de España.

———. 2011. "Optimal Research and Development and the Cost of Business Cycles." *Journal of Economic Growth* 16 (3) (April 26): 257–283.

Pavlov, Oscar, and Mark Weder. 2012. "Variety Matters." *Journal of Economic Dynamics and Control* 36 (4) (April): 629–641.

Peretto, Pietro F. 1999. "Cost Reduction, Entry, and the Interdependence of Market Structure and Economic Growth." *Journal of Monetary Economics* 43 (1) (February 19): 173–195.

Phillips, Kerk L., and Jeff Wrase. 2006. "Is Schumpeterian `creative Destruction' a Plausible Source of Endogenous Real Business Cycle Shocks?" *Journal of Economic Dynamics and Control* 30 (11) (November): 1885–1913.

Pincus, Laura B. 1991. "Computation of Damages in Patent Infringement Actions." *Harvard Journal of Law & Technology* 5: 95.

Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50 (1) (January): 97–109.

Rudebusch, Glenn D., and Eric T. Swanson. 2012. "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks." *American Economic Journal: Macroeconomics* 4 (1): 105–43.

Scott, John. 1984. "Firm versus Industry Variability in R&D Intensity." In *R & D, Patents, and Productivity*. NBER.

Serrano, C. J. 2007. "Technology Transfer and the Business Cycle: Evidence from Patent Reassignments."

Serrano, Carlos J. 2010. "The Dynamics of the Transfer and Renewal of Patents." *The RAND Journal of Economics* 41 (4) (December 1): 686–708.

Siliverstovs, Boriss, and D'Artis Kancs. 2012. *R&D and Non-Linear Productivity Growth of Heterogeneous Firms*. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

The Associated Press. 2008. "Court Rules Business Concept Cannot Be Patented." *The New York Times*, October 31, sec. Business.

Tingvall, Patrik Gustavsson, and Andreas Poldahl. 2006. "Is There Really an Inverted U-Shaped Relation between Competition and R&D?" *Economics of Innovation and New Technology* 15 (3): 101.

Wälde, Klaus. 2005. "Endogenous Growth Cycles." *International Economic Review* 46 (3): 867–894.

7. Online appendices

7.1. The free-entry and first order conditions

When deciding how much research and appropriation to perform, firms realise that if they perform a non-equilibrium amount then in the next period they will have an incentive to set a different mark-up to the other firms in their industry. (The clearest example of this is when we have perfect competition, in which case the most productive firm would want to price just below the second most productive firms' marginal cost.) It may be seen that in non-symmetric equilibrium the optimal price satisfies:

$$P_t(i,j) = \frac{W_t^{\mathrm{W}}}{A_t(i,j)}\left[1 + \frac{\eta\lambda}{1-(1-\eta)\frac{1}{J_{t-1}(i)}\left(\frac{P_t(i,j)}{P_t(i)}\right)^{-\frac{1}{\eta\lambda}}}\right].$$

Since we are looking for a symmetric equilibrium, it is sufficient to approximate this locally around $P_t(i) = P_t(i,j)$ in order to calculate firms' research and appropriation incentives. Taking a log-linear approximation of $\log P_t(i,j)$ in $\frac{P_t(i,j)}{P_t(i)}$ gives us that:

$$P_t(i,j) \approx \frac{W_t^{\mathrm{W}}}{A_t(i,j)}\big(1+\mu_{t-1}(i)\big)\left(\frac{P_t(i,j)}{P_t(i)}\right)^{-\omega_{t-1}(i)}$$

where $\omega_t(i) := \frac{J_t(i)(1-\eta)}{\left(J_t(i)-(1-\eta)\right)^2(1+\mu_t(i))}$ captures the strength of these incentives to deviate from setting the same mark-up as all other firms in their industry. Therefore $P_t(i) \approx \frac{W_t^{\mathrm{W}}}{A_t(i)}\big(1+\mu_{t-1}(i)\big)$ and $P_t(i,j) \approx \frac{W_t^{\mathrm{W}}}{A_t(i,j)}\big(1+\mu_{t-1}(i)\big)\left(\frac{A_t(i,j)}{A_t(i)}\right)^{\frac{\omega_{t-1}(i)}{1+\omega_{t-1}(i)}}$ where:

$$A_t(i) := \left[\frac{1}{J_{t-1}(i)}\sum_{j=1}^{J_{t-1}(i)} A_t(i,j)^{\frac{1}{\eta\lambda(1+\omega_{t-1}(i))}}\right]^{\eta\lambda(1+\omega_{t-1}(i))}.$$

Therefore, up to a first order approximation around the symmetric solution, profits are given by:

$$\beta\frac{1}{|\mathbb{I}_t|J_t(i)}\left(\frac{1+\mu_t}{1+\mu_t(i)}\right)^{\frac{1}{\lambda}}\mathbb{E}_t\Xi_{t+1}X_{t+1}\left[\left(\frac{A_{t+1}(i,j)}{A_{t+1}(i)}\right)^{\frac{\omega_t(i)}{1+\omega_t(i)}} - \frac{1}{1+\mu_t(i)}\right]\left(\frac{A_{t+1}(i,j)}{A_{t+1}(i)}\right)^{\frac{1-\eta\lambda\omega_t(i)}{\eta\lambda(1+\omega_t(i))}}\left(\frac{A_{t+1}(i)}{A_{t+1}}\right)^{\frac{1}{\lambda}}$$
$$-\left[L_t^{\mathrm{R}}(i,j) + L_t^{\mathrm{A}}(i,j) + L_t^{\mathcal{R}}(i) + L^{\mathrm{F}}\right]W_t^{\mathrm{W}}.$$

Note that if $J_t(i) > \frac{2\sqrt{2}(3-\sqrt{2})}{1+2\sqrt{2}} \approx 1.17$, then $1-\eta\lambda\omega_t(i) > 0$ (by tedious algebra), so providing there are at least two firms in the industry, this expression is guaranteed to be increasing and concave in $A_{t+1}(i,j)$.

Let $m_t^{\mathrm{R}}(i,j)W_t^{\mathrm{W}}$ be the Lagrange multiplier on research's positivity constraint and $m_t^{\mathrm{A}}(i,j)W_t^{\mathrm{W}}$ be the Lagrange multiplier on appropriation's positivity constraint. Then in a symmetric equilibrium, the two first order conditions and the free entry condition (respectively) mean:

$$\beta\frac{1}{|\mathbb{I}_t|J_t(i)}\frac{\mu_t(i)}{1+\mu_t(i)}\left(\frac{1+\mu_t}{1+\mu_t(i)}\right)^{\frac{1}{\lambda}}\mathbb{E}_t\Xi_{t+1}X_{t+1}\left(\frac{A_{t+1}(i)}{A_{t+1}}\right)^{\frac{1}{\lambda}}\frac{d_t(i)}{\mu_t(i)}\frac{Z_{t+1}E_t^{\mathrm{R}}(i)}{1+\gamma Z_{t+1}E_t^{\mathrm{R}}(i)L_t^{\mathrm{R}}(i)} = W_t^{\mathrm{W}}\big(1-m_t^{\mathrm{R}}(i)\big)$$

$$\beta\frac{1}{|\mathbb{I}_t|J_t(i)}\frac{\mu_t(i)}{1+\mu_t(i)}\left(\frac{1+\mu_t}{1+\mu_t(i)}\right)^{\frac{1}{\lambda}}\mathbb{E}_t\Xi_{t+1}X_{t+1}\left(\frac{A_{t+1}(i)}{A_{t+1}}\right)^{\frac{1}{\lambda}}\frac{d_t(i)}{\mu_t(i)}\frac{1+(\gamma-\zeta^{\mathrm{R1}})Z_{t+1}E_t^{\mathrm{R}}(i)L_t^{\mathrm{R}}(i)}{1+\gamma Z_{t+1}E_t^{\mathrm{R}}(i)L_t^{\mathrm{R}}(i)}$$
$$\cdot\frac{1}{\tau}\frac{E_t^{\mathrm{A}}(i)\left(A_t^{*\tau} - A_t(i)^{\tau}\right)}{A_t^{**}(i)^{\tau}\left(1+E_t^{\mathrm{A}}(i)L_t^{\mathrm{A}}(i)\right)^2} = W_t^{\mathrm{W}}\big(1-m_t^{\mathrm{A}}(i)\big)$$

$$\beta\frac{1}{|\mathbb{I}_t|J_t(i)}\frac{\mu_t(i)}{1+\mu_t(i)}\left(\frac{1+\mu_t}{1+\mu_t(i)}\right)^{\frac{1}{\lambda}}\mathbb{E}_t\Xi_{t+1}X_{t+1}\left(\frac{A_{t+1}(i)}{A_{t+1}}\right)^{\frac{1}{\lambda}} = \left[L_t^{\mathrm{R}}(i,j) + L_t^{\mathrm{A}}(i,j) + L_t^{\mathcal{R}}(i) + L^{\mathrm{F}}\right]W_t^{\mathrm{W}}$$

where:

$$d_t(i) := 1 - \frac{\omega_t(i)}{1+\omega_t(i)}\frac{\big(\lambda-\mu_t(i)\big)\big(\mu_t(i)-\eta\lambda\big)}{\lambda(1-\eta)\mu_t(i)} < 1$$

and where we have dropped j indices on variables which are the same across the industry.

We also have that:

$$\frac{\big(\lambda-\mu_t(i)\big)\big(\mu_t(i)-\eta\lambda\big)}{\lambda(1-\eta)\mu_t(i)}\le\frac{\big(1-\sqrt{\eta}\big)\big(\sqrt{\eta}-\eta\big)}{(1-\eta)\sqrt{\eta}}=\frac{1-\sqrt{\eta}}{1+\sqrt{\eta}}<1$$

so $d_t(i)>0$, as $\omega_t(i)<\frac{1}{\eta\lambda}$. In fact, we may derive tighter bounds on $d_t(i)$. Tedious algebra (available on request) gives that as $J_t(i)\ge 2$, $d_t(i)>\frac{2}{3}$, with equality in the limit as $\eta\to 0$ and $J_t(i)\to 2$.

That the solution for research when $Z_{t+1}\equiv 1$ is given by equation (1.3) is a trivial consequence of the complementary slackness condition and the facts that $\frac{1}{\mu_t(i)}<\gamma$ and $d_t(i)<1$. Deriving (1.4) is less trivial though.

Begin by defining $k_t(i):=\frac{1+(\gamma-\zeta^{\mathrm{R1}})\mathcal{L}_t^{\mathrm{R}}(i)}{1+\gamma\mathcal{L}_t^{\mathrm{R}}(i)}$, and note that since we are assuming $\gamma>\zeta^{\mathrm{R1}}\ge 0$, we have that $0<k_t(i)\le 1$. Also define:

$$n_t(i):=\frac{d_t(i)k_t(i)}{\tau\mu_t(i)}E_t^{\mathrm{A}}(i)\left[\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}-1\right]\left[L_t^{\mathrm{R}}(i)+L_t^{\mathcal{R}}(i)+L^{\mathrm{F}}\right]\ge 0,$$

which is not a function of $L_t^{\mathrm{A}}(i)$, given $L_t^{\mathrm{R}}(i)$.

We can then combine the appropriation first order condition with the free entry condition to obtain:

$$\frac{1}{\big(1+\mathcal{L}_t^{\mathrm{A}}(i)\big)^2}\left(\frac{A_t^*(i)}{A_t^{**}(i)}\right)^{\tau}\left[\frac{d_t(i)k_t(i)}{\tau\mu_t(i)}\left[\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}-1\right]\mathcal{L}_t^{\mathrm{A}}(i)+n_t(i)\right]=1-m_t^{\mathrm{A}}(i).$$

Since the left hand side is weakly positive, from the dual feasibility condition we know $m_t^{\mathrm{A}}(i)\in[0{,}1]$. Now when $L_t^{\mathrm{A}}(i)=0$, this becomes:

$$n_t(i)=1-m_t^{\mathrm{A}}(i),$$

since in this case $A_t^*(i)=A_t^{**}(i)$. Therefore when $L_t^{\mathrm{A}}(i)=0$, $n_t(i)\le 1$.

We now prove the converse. Suppose then for a contradiction that $L_t^{\mathrm{A}}(i)>0$, but $n_t(i)\le 1$. By complementary slackness, we must have $m_t^{\mathrm{A}}(i)=0$, hence:

$$\begin{aligned}1\ge n_t(i)&=\big(1+\mathcal{L}_t^{\mathrm{A}}(i)\big)^2\left(\frac{A_t^{**}(i)}{A_t^*(i)}\right)^{\tau}-\frac{d_t(i)k_t(i)}{\tau\mu_t(i)}\left[\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}-1\right]\mathcal{L}_t^{\mathrm{A}}(i)\\&\ge\big(1+\mathcal{L}_t^{\mathrm{A}}(i)\big)^2\left(\frac{A_t^{**}(i)}{A_t^*(i)}\right)^{\tau}-\left[\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}-1\right]\mathcal{L}_t^{\mathrm{A}}(i)\\&=\big(1+\mathcal{L}_t^{\mathrm{A}}(i)\big)\left[\big(1+\mathcal{L}_t^{\mathrm{A}}(i)\big)+\mathcal{L}_t^{\mathrm{A}}(i)\left[\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}-1\right]\right]-\left[\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}-1\right]\mathcal{L}_t^{\mathrm{A}}(i),\end{aligned}$$

where we have used the facts that $d_t(i)k_t(i)\le 1$ and $\frac{1}{\mu_t(i)}<\tau$ to derive the second inequality.

Expanding the brackets then gives that:

$$1\ge 1+2\mathcal{L}_t^{\mathrm{A}}(i)+\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}\mathcal{L}_t^{\mathrm{A}}(i)^2,$$

i.e. that $0\ge 2+\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}\mathcal{L}_t^{\mathrm{A}}(i)$ which is a contradiction as $\left(\frac{A_t^*}{A_t^*(i)}\right)^{\tau}\mathcal{L}_t^{\mathrm{A}}(i)\ge 0$.

We have proven then that providing $\frac{1}{\mu_t(i)}<\tau$, $L_t^{\mathrm{A}}(i)=0$ if and only if $n_t(i)\le 1$. It just remains for us to solve for $L_t^{\mathrm{A}}(i)$ when it is strictly positive. From the above, we have that, in this case:

$$\left(\frac{A_t^*(i)}{A_t^*}\right)^{\tau}[n_t(i)-1]=2\left[1-\frac{1}{2}\left[1+\frac{d_t(i)k_t(i)}{\tau\mu_t(i)}\right]\left[1-\left(\frac{A_t^*(i)}{A_t^*}\right)^{\tau}\right]\right]\mathcal{L}_t^{\mathrm{A}}(i)+\mathcal{L}_t^{\mathrm{A}}(i)^2.$$

Hence:

$$\mathcal{L}_t^{\mathrm{A}}(i) = -\left[1 - \frac{1}{2}\left[1 + \frac{d_t(i)k_t(i)}{\tau\mu_t(i)}\right]\left[1 - \left(\frac{A_t^*(i)}{A_t^*}\right)^\tau\right]\right] + \sqrt{\left[1 - \frac{1}{2}\left[1 + \frac{d_t(i)k_t(i)}{\tau\mu_t(i)}\right]\left[1 - \left(\frac{A_t^*(i)}{A_t^*}\right)^\tau\right]\right]^2 + \left(\frac{A_t^*(i)}{A_t^*}\right)^\tau [n_t(i) - 1]},$$

since the lower solution is guaranteed to be negative as $n_t(i) > 1$ when $L_t^{\mathrm{A}}(i) > 0$.

7.2. The steady state for non-patent-protected industries

In an industry i which is not patent-protected and in which appropriation, but no research, is performed, from (1.1) and (1.4), we have that:

$$f_t(i) + \sqrt{f_t(i)^2 + g_t(i)} = \mathcal{L}_t^{\mathrm{A}}(i) = \left[1 - \frac{\left(\frac{A_{t+1}^*(i)}{A_t^*(i)}\right)^\tau - 1}{1 - \left(\frac{A_t^*(i)}{A_t^*}\right)^\tau}\left(\frac{A_t^*(i)}{A_t^*}\right)^\tau\right]^{-1} - 1.$$

If we treat $\mathfrak{p}_1 := \tau\frac{\mu_t(i)}{d_t(i)} - 1 \approx 0$, $\mathfrak{p}_2 := E_t^{\mathrm{A}}(i)L^{\mathrm{F}} \approx 0$ and $\mathfrak{p}_3 := \left(\frac{A_{t+1}^*(i)}{A_t^*(i)}\right)^\tau - 1 \approx 0$ as fixed, this leaves us with a cubic in $\left(\frac{A_t^*(i)}{A_t^*}\right)^\tau$, for which only one solution will be feasible (i.e. strictly less than 1). Taking a second order Taylor approximation of this solution in $\mathfrak{p}_1$, $\mathfrak{p}_2$ and $\mathfrak{p}_3$, reveals (after some messy computation), that:

$$\left(\frac{A_t^*(i)}{A_t^*}\right)^\tau \approx \mathfrak{p}_2\big(1 - (\mathfrak{p}_1 + \mathfrak{p}_2)\big) = E_t^{\mathrm{A}}(i)L_t^{\mathrm{F}}\left(2 - \tau\frac{\mu_t(i)}{d_t(i)} - E_t^{\mathrm{A}}(i)L^{\mathrm{F}}\right)$$

(The effect of $\mathfrak{p}_3$ on $\left(\frac{A_t^*(i)}{A_t^*}\right)^\tau$ is third order and hence it does not appear in this expression.)

From this approximate solution for $\left(\frac{A_t^*(i)}{A_t^*}\right)^\tau$ then, we have that the relative productivity of a non-protected industry is decreasing in its mark-up. Furthermore, from dropping to a first order approximation, we have that $A_t^*(i)^{1+\frac{\zeta^{\mathrm{A1}}}{\tau}} \approx A_t^{*1+\frac{\zeta^{\mathrm{A2}}}{\tau}}\left(\mathbb{I}_t^{\phi^{\mathrm{A}}}\Psi^{\mathrm{A}}L^{\mathrm{F}}\right)^{\frac{1}{\tau}}$, so asymptotically non-protected industries are growing at $\left[1 + \frac{\zeta^{\mathrm{A2}}}{\tau}\right]\left[1 + \frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1}$ times the growth rate of the frontier, plus $\frac{\phi^{\mathrm{A}}}{\tau}\left[1 + \frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1}$ times the growth rate of the measure of industries.

To examine the long run behaviour of this level of relative productivity, recall that we assumed $A_t^{*-\zeta^{\mathrm{A}}}\mathbb{I}_t^{\phi^{\mathrm{A}}} \to 0$ as $t \to \infty$, where $\zeta^{\mathrm{A}} = \zeta^{\mathrm{A1}} - \zeta^{\mathrm{A2}} > 0$. Thus if we write $g_{\mathcal{V}}$ for the asymptotic growth rate of the variable $\mathcal{V}$, we have $\phi^{\mathrm{A}}g_{\mathbb{I}} < \zeta^{\mathrm{A1}}g_{A^*} - \zeta^{\mathrm{A2}}g_{A^*}$, so:

$$\begin{aligned} g_{A^*(i)} &= \left[1 + \frac{\zeta^{\mathrm{A2}}}{\tau}\right]\left[1 + \frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1} g_{A^*} + \frac{\phi^{\mathrm{A}}}{\tau}\left[1 + \frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1} g_{\mathbb{I}} \\ &< \left[1 + \frac{\zeta^{\mathrm{A2}}}{\tau}\right]\left[1 + \frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1} g_{A^*} + \frac{\zeta^{\mathrm{A1}}}{\tau}\left[1 + \frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1} g_{A^*} - \frac{\zeta^{\mathrm{A2}}}{\tau}\left[1 + \frac{\zeta^{\mathrm{A1}}}{\tau}\right]^{-1} g_{A^*} = g_{A^*}. \end{aligned}$$

7.3. The inventor-firm bargaining process

We model the entire process of setting and paying rents as follows:

1) Firms enter, paying the fixed cost.
2) Firms who have entered conduct appropriation, then research.
3) The "idea shock" for next period's production, Z_{t+1}, is realised and firms and patent holders learn its level.

4) Finally, firms arrive at the patent-holder to conduct bargaining, with these arrivals taking place sequentially but in a random order. (For example, all firms phone the patent-holder sometime in the week before production is to begin.) In this bargaining we suppose that the patent-holder has greater bargaining power, since they have a longer outlook[40] and since they lose nothing if bargaining collapses[41]. We also suppose that neither patent-holders nor firms are able to observe or verify either how many (other) firms paid the fixed cost, or what research and appropriation levels they chose. This is plausible because until production begins it is relatively easy to keep such things hidden (for example, by purchasing the licence under a spin-off company), and because it is hard to ascertain ahead of production exactly what product a firm will be producing. We assume bargaining takes an alternating offer form, (Rubinstein 1982) but that it happens arbitrarily quickly (i.e. in the no discounting limit).
5) Firms pay the agreed rents if bargaining was successful. Since this cost is expended before production, we continue to suppose firms have to borrow in the period before production in order to cover it. Firms will treat it as a fixed cost, sunk upon entry, since our unobservability assumptions mean bargaining's outcome will not be a function of research and appropriation levels.
6) The next period starts, other aggregate shocks are realised and production takes place.
7) The patent-holder brings court cases against any firms who produced but decided not to pay the rent. For simplicity, we assume the court always orders the violating firm to pay damages to the patent-holder, which are given as follows:
   a) When the courts believe rents were not reasonable (i.e. $L_t^{\mathcal{R}}(i) > L_t^{\mathcal{R}*}(i)$, where $L_t^{\mathcal{R}*}(i)W_t^{\mathrm{W}}$ is the level courts determine to be "reasonable royalties"), they set damages greater than $L_t^{\mathcal{R}*}(i)W_t^{\mathrm{W}}$, as "*the infringer would have nothing to lose, and everything to gain if he could count on paying only the normal, routine royalty non-infringers might have paid*"[42]. We assume excess damages over $L_t^{\mathcal{R}*}(i)W_t^{\mathrm{W}}$ are less than the patent-holder's legal costs however.
   b) When the courts consider the charged rent to have been reasonable (i.e. $L_t^{\mathcal{R}}(i) \le L_t^{\mathcal{R}*}(i)$) the courts award punitive damages of more than $\max\left\{ L_t^{\mathcal{R}*}(i)W_t^{\mathrm{W}}, \left(\frac{1}{1-\wp}\right) L_t^{\mathcal{R}}(i)W_t^{\mathrm{W}}, \right\}$, where $\wp$ is the bargaining power of the firm, in the sense of the generalized Nash bargaining solution.[43]

Under this specification:

$$L_t^{\mathcal{R}}(i) = \min\{L_t^{\mathcal{R}*}(i), (1-\wp)\left[L_t^{\mathrm{R}}(i) + L_t^{\mathrm{A}}(i) + L_t^{\mathcal{R}}(i) + L^{\mathrm{F}}\right]\}$$

since entry is fixed when bargaining takes place, since patent-holders know that bargaining to a rent level any higher than $L_t^{\mathcal{R}*}(i)W_t^{\mathrm{W}}$ will just result in them having to pay legal costs,[44] and since $\left[L_t^{\mathrm{R}}(i) + L_t^{\mathrm{A}}(i) + L_t^{\mathcal{R}}(i) + L^{\mathrm{F}}\right]W_t^{\mathrm{W}}$ is equal to the production period profits of each firm in industry i, by the free entry condition.[45] Therefore, in equilibrium:

[40] Consider what happens as the time gap between offers increases. When this gap is large enough only one offer would be made per-period, meaning the patent-holder would make a take-it-or-leave-it offer giving (almost) nothing to the firm, which the firm would then accept.
[41] The firm owner may, for example, face restrictions from starting businesses in future if as a result of the bargaining collapse they are unable to repay their creditors.
[42] Panduit Corp. v. Stahlin Brothers Fibre Works, Inc., 575 F.2d 1152, 1158 (6th Circuit 1978), cited in Pincus (1991).
[43] The level $\left(\frac{1}{1-\wp}\right)L_t^{\mathcal{R}}(i)W_t^{\mathrm{W}}$ is chosen to ensure that, with equilibrium rents, firms prefer not to produce at all rather than to produce without paying rents.
[44] The disagreement point is zero since it is guaranteed that $L_t^{\mathfrak{R}}(i) \le L_t^{\mathfrak{R}*}(i)$ and so punitive damages would be awarded were the firm to produce without paying rents, which, by construction, leaves them worse off than not producing.
[45] A similar expression can also be derived if we assume instead that courts guarantee infringers a fraction $\wp$ of production profits, or if we assume courts always award punitive damages but firms are able to hide a fraction $\wp$ of their production profits.

$$L_t^{\mathcal{R}}(i) = \min\{L_t^{\mathcal{R}*}(i), L_t^{\mathcal{R}\dagger}(i)\}, \tag{5.1}$$

where $L_t^{\mathcal{R}\dagger}(i)$ is a solution to equations (1.3), (1.4) and (1.5), (i.e. $L_t^{\mathcal{R}}(i) = \frac{1-\rho}{\rho}\left[L_t^{\mathrm{R}}(i) + L_t^{\mathrm{A}}(i) + L^{\mathrm{F}}\right]$) if one exists, or $+\infty$ otherwise. Because damages are always greater than $L_t^{\mathcal{R}*}(i)W_t^{\mathrm{W}}$, these rents will be sufficiently low to ensure firms are always prepared to licence the patent at the bargained price in equilibrium.

Now suppose we are out of equilibrium and fewer firms than expected have entered. Since neither the patent-holder nor firms can observe how many firms have entered, and since firms arrive at the patent-holder sequentially, both sides will continue to believe that the equilibrium number of firms has entered and so rents will not adjust. On the other hand, suppose that (out of equilibrium) too many firms enter. When the first unexpected firm arrives at the patent-holder to negotiate, the patent-holder will indeed realise that too many firms have entered. However, since the firm they are bargaining with has no way of knowing this,[46] the patent-holder can bargain for the same rents as in equilibrium. Therefore, even out of equilibrium:

$$L_t^{\mathcal{R}}(i) = \min\{L_t^{\mathcal{R}*}(i), L_t^{\mathcal{R}\dagger}(i)\}$$

where we stress $L_t^{\mathcal{R}\dagger}(i)$ is not a function of the decisions any firm happened to take. This ensures that any solution of equation (5.1) with equations (1.3) and (1.4), will also be an equilibrium, even allowing for the additional condition that the derivative of firm profits with respect to the number of firms must be negative at an optimum.

We now just have to pin down "reasonable royalties", $L_t^{\mathcal{R}*}(i)W_t^{\mathrm{W}}$ *Georgia-Pacific*, 318 F. Supp. at 1120 (S.D.N.Y. 1970), *modified on other grounds*, 446 F.2d 295 (2d Cir.), *cert. denied*, 404 U.S. 870 (1971), cited in Pincus (1991), defines a reasonable royalty as "*the amount that a licensor (such as the patentee) and a licensee (such as the infringer) would have agreed upon (at the time the infringement began) if both had been reasonably and voluntarily trying to reach an agreement; that is, the amount which a prudent licensee—who desired, as a business proposition, to obtain the licence to manufacture and sell a particular article embodying the patented invention—would have been willing to pay as a royalty and yet be able to make a reasonable profit and which amount would have been acceptable by a prudent patentee who was willing to grant a licence.*"

Certainly it must be the case that $L_t^{\mathcal{R}*}(i) \le L_t^{\overline{\mathcal{R}}}(i)$, where $L_t^{\overline{\mathcal{R}}}(i)$ is the level of rents at which $J_t(i) = 1$, since rents so high that no one is prepared to pay them must fall foul of the courts' desire to ensure licensees can make a profit.[47] However, since when $J_t(i) = 1$ the sole entering firm (almost) may as well be the patent-holder themselves, where possible the courts will set $L_t^{\mathcal{R}*}(i)$ sufficiently low to ensure that $J_t(i) > 1$ in equilibrium, again following the idea that licensees ought to be able to make a profit. When there is a $J_t(i) > 1$ solution to equations (1.3), (1.4) and (1.5) already (i.e. $L_t^{\mathcal{R}\dagger}(i) < \infty$), the courts will just set $L_t^{\mathcal{R}*}(i)$ at the rent level that would obtain in that solution, thus preventing the possibility of $J_t(i) = 1$ being an equilibrium. It may be shown that for sufficiently large t such a solution is guaranteed to exist, so in this case $L_t^{\mathcal{R}*}(i) = L_t^{\mathcal{R}\dagger}(i) = L_t^{\mathcal{R}}(i)$.[48]

[46] Either they are a firm that thinks the equilibrium number of firms has entered, or they are a firm that thinks more than the equilibrium number of firms has entered, but that does not know whether the patent-holder has yet realised this.

[47] "*…the very definition of a reasonable royalty assumes that, after payment, the infringer will be left with a profit.*" Georgia-Pacific Corp. v. U.S. Plywood-Champion Papers Corp., 446 F.2d 295, 299 & n.1 (2d Cir.), *cert. denied*, 404 U.S. 870 (1971), cited in Pincus (1991).

[48] There may still be multiple solutions for rents (as (1.2), (1.3) and (1.4) from Holden (2013) might have multiple solutions), but of these only the one with minimal entry is really plausible, since this is both weakly Pareto dominant (firms always make zero profits and it may be shown that the patent-holder prefers minimal entry) and less risky for entering firms (if entering firms are unsure if the patent-holder will play the high rent or the low rent equilibrium, they are always better off assuming the high rent one since if that assumption is wrong they make strict profits, whereas had they assumed low rents but rents were in fact high they would make a strict loss).

7.4. Proof of instantaneous catch-up in frontier industries

Suppose $(i_t)_{t=0}^{\infty}$ is a sequence of industries, all protected at t, for which the sequence $\left(A_t^*(i_t)\right)_{t=0}^{\infty}$ grows at rate $\tilde{g} \le g_{A^*}$ asymptotically. We seek to prove:

(a) that $\lim_{t\to\infty} A_t^{**}(i_t)^{-\zeta^{R1}} A_t^{*\zeta^{R2}} |\mathbb{I}_t|^{\phi^R} \Psi^R L_t^A(i_t) = 0$, and

(b) that the existence of $\varepsilon > 0$ such that $\frac{A_t^*(i_t)}{A_t^*} < 1 - \varepsilon$ for all sufficiently large t implies $\lim_{t\to\infty} \mathcal{L}_t^A(i) = \infty$.

First, note that since mark-ups are asymptotically bounded above and below, so too is effective research. Consequently, $\not{f}_t(i)$ is bounded above, so from (1.4) we know that the asymptotic growth rate of $\mathcal{L}_t^A(i_t)$ is equal to that of $\sqrt{g_t(i_t)}$. Hence, if the antecedent of (b) holds, the asymptotic growth rate of $\mathcal{L}_t^A(i_t)$ is bounded below by:

$$\frac{1}{2}\left[(\zeta^{R1} - \zeta^{A1})\tilde{g} + (\zeta^{A2} - \zeta^{R2})g_{A^*} + (\phi^A - \phi^R)g_{\mathbb{I}}\right] > 0$$

(as $\zeta^{R1} > \zeta^{A1}$, $\zeta^{R2} \le \zeta^{A2}$ and $\phi^R \le \phi^A$), which establishes the consequent of (b).

In any case, the asymptotic growth rate of $\mathcal{L}_t^A(i_t)$ is less or equal to:

$$\frac{1}{2}\left[(\zeta^{R1}g_{A^*} - \zeta^{A1}\tilde{g}) + (\zeta^{A2} - \zeta^{R2})g_{A^*} + (\phi^A - \phi^R)g_{\mathbb{I}}\right].$$

Therefore the asymptotic growth rate of $A_t^{**}(i_t)^{-\zeta^{R1}} A_t^{*\zeta^{R2}} \mathbb{I}_t^{\phi^R} \Psi^R L_t^A(i_t)$ is less or equal to

$$-\zeta^{R1}\tilde{g} + \zeta^{A1}\tilde{g} + \frac{1}{2}\left[(\zeta^{R1}g_{A^*} - \zeta^{A1}\tilde{g}) - (\zeta^{A2} - \zeta^{R2})g_{A^*} - (\phi^A - \phi^R)g_{\mathbb{I}}\right].$$

For (a) to be proven, we then just need that $\frac{\zeta^{R1}}{2\zeta^{R1} - \zeta^{A1}} g_{A^*} < \tilde{g}$.

We require (a) and (b) to hold for any sequence of protected industries. Conditions for this may be derived using the fact that for any such sequence, $g_S + g_{A^*} \le \tilde{g} \le g_{A^*}$. Ensuring that $\frac{\zeta^{R1}}{2\zeta^{R1} - \zeta^{A1}} g_{A^*} < \tilde{g}$ for any $\tilde{g}$ in this interval is equivalent to requiring that $\frac{\zeta^{R1}}{2\zeta^{R1} - \zeta^{A1}} g_{A^*} < g_S + g_{A^*}$. I.e. we require that $-g_S < \frac{\zeta^{R1} - \zeta^{A1}}{2\zeta^{R1} - \zeta^{A1}} g_{A^*}$. Since the right hand side is strictly positive (as $\zeta^{R1} > \zeta^{A1}$) this holds for a positive measure of values for g_S.

7.5. The de-trended model

Below we give the equations of the stationary model to which the model described in section 3 converges as $t \to \infty$.

7.5.1. Households

- **Stochastic discount factor:** $\Xi_t = \frac{\hat{C}_{t-1}}{\hat{C}_t G_{A,t}}$, where $\hat{C}_t := \frac{C_t}{N_t A_t}$ is consumption per person in labour supply units and $G_{\mho,t} = \frac{\mho_t}{\mho_{t-1}}$.
- **Labour supply:** $\hat{L}_t^{S^\nu} = \frac{\widehat{W}_t^W}{\hat{C}_t}$, where $\hat{L}_t^S := \frac{L_t^S}{N_t}$ is labour supply per person and $\widehat{W}_t^W := \frac{W_t^W}{A_t}$ is the wage per effective unit of labour supply.
- **Euler equation:** $\beta R_t \mathbb{E}_t[\Xi_{t+1}] = 1$, where R_t is the real interest rate.

7.5.2. Aggregate relationships

- **Aggregate mark-up pricing:** $\widehat{W}_t^W = \frac{1}{1+\mu_{t-1}}$ where μ_{t-1} is the aggregate mark-up in period t.

- **Mark-up aggregation:** $\left(\frac{1}{1+\mu_t}\right)^{\frac{1}{\lambda}} = \left(\frac{1}{1+\mu_t^{\mathrm{P}}}\right)^{\frac{1}{\lambda}} \eth_t + \left(\frac{1}{1+\eta\lambda}\right)^{\frac{1}{\lambda}}(1-\eth_t)$, where $\mu_t^{\mathrm{P}} = \mu_t(\sup \mathbb{I}_t)$ is the mark-up in any protected industry at $t+1$, and $\eth_t := 1 - \left[1-(1-q)\eth_{t-1}\right]\frac{1-\delta_{\mathbb{I},t}}{G_{|\mathbb{I}|,t}}$ is the proportion of industries that will produce a patent protected product in period $t+1$.
- **Productivity aggregation:** $\left(\frac{\hat{A}_t}{1+\mu_{t-1}}\right)^{\frac{1}{\lambda}} = \left(\frac{1}{1+\mu_{t-1}^{\mathrm{P}}}\right)^{\frac{1}{\lambda}} \eth_{t-1} + \left(\frac{\hat{A}_t^{\mathrm{N}}}{1+\eta\lambda}\right)^{\frac{1}{\lambda}}(1-\eth_{t-1})$, where $\hat{A}_t := \frac{A_t}{A_t^*}$ is aggregate productivity relative to the frontier[49] and $\hat{A}_t^{\mathrm{N}} := \left[\left(\frac{1}{G_{A^*,t}}\right)^{\frac{1}{\lambda}}\left(\frac{q}{1/\eth_{t-2}-(1-q)}\right) + \left(\frac{\hat{A}_{t-1}^{\mathrm{N}}}{G_{A^*,t}}\right)^{\frac{1}{\lambda}}\left(1-\frac{q}{1/\eth_{t-2}-(1-q)}\right)\right]^{\lambda}$ is the aggregate relative productivity of non-protected industries.

7.5.3. Firm decisions

- **Strategic in-industry pricing:** $\mu_t^{\mathrm{P}} = \lambda \frac{\eta J_t^{\mathrm{P}}}{J_t^{\mathrm{P}}-(1-\eta)}$, where $J_t^{\mathrm{P}} := J_t(\sup \mathbb{I}_t)$ is the number of firms in a protected industry performing research at t.
- **Firm research decisions:** $\frac{d_t^{\mathrm{P}}}{p\mu_t^{\mathrm{P}}}\mathbb{E}_t \Xi_{t+1} G_{X,t+1} \hat{A}_{t+1}^{-\frac{1}{\lambda}} \frac{Z_{t+1}\mathcal{L}_t^{\mathrm{RP}}}{1+\gamma Z_{t+1}\mathcal{L}_t^{\mathrm{RP}}} = \mathbb{E}_t \Xi_{t+1} G_{X,t+1}\hat{A}_{t+1}^{-\frac{1}{\lambda}}$, where $\mathcal{L}_t^{\mathrm{RP}} := \mathcal{L}_t^{\mathrm{R}}(\sup \mathbb{I}_t)$ is the amount of effective research conducted by firms in protected industries $d_t^{\mathrm{P}} := d_t(\sup \mathbb{I}_t)$ is the value of $d_t(i)$ in these industries. (This equation means that $\mathcal{L}_t^{\mathrm{RP}} \approx \frac{p\mu_t^{\mathrm{P}}}{d_t^{\mathrm{P}}-p\gamma\mu_t^{\mathrm{P}}}$.)
- **Research and appropriation payoff:** $G_{A^*,t} = \left(1+\gamma Z_t \mathcal{L}_{t-1}^{\mathrm{RP}}\right)^{\frac{1}{\gamma}}$.
- **Free entry of firms:** $\beta \frac{1}{\hat{\mathbb{I}}_t J_t^{\mathrm{P}}} \frac{\mu_t^{\mathrm{P}}}{1+\mu_t^{\mathrm{P}}}\left(\frac{1+\mu_t}{1+\mu_t^{\mathrm{P}}}\right)^{\frac{1}{\lambda}} \mathbb{E}_t \Xi_{t+1} G_{X,t+1} \hat{A}_{t+1}^{-\frac{1}{\lambda}} = \frac{1}{p}\mathcal{L}_t^{\mathrm{RP}} \frac{\widehat{W}_t^{\mathrm{W}}}{\widehat{X}_t}$, where $\hat{\mathbb{I}}_t := \frac{|\mathbb{I}_t|}{N_t E_t^{\mathrm{R}}(\sup \mathbb{I}_t)}$ is the measure of products relative to its trend,[50] and $\widehat{X}_t := \frac{X_t}{N_t A_t}$ is output per person in labour supply units.

7.5.4. Inventor decisions

- **Inventor value function:** is given recursively by:
$\widehat{V}_t^{\mathbb{I}} = \frac{1-p}{p}\mathcal{L}_t^{\mathrm{RP}}\widehat{W}_t^{\mathrm{W}} J_t^{\mathrm{P}} + \beta(1-q)\mathbb{E}_t(1-\delta_{\mathbb{I},t+1})\Xi_{t+1}\frac{G_{A,t+1}}{G_{E^{\mathrm{R}},t+1}}\widehat{V}_{t+1}^{\mathbb{I}}$, where $\widehat{V}_t^{\mathbb{I}} := \frac{V_t^{\mathbb{I}}(\sup \mathbb{I}_t)E_t^{\mathrm{R}}(\sup \mathbb{I}_t)}{A_t}$ and $G_{E^{\mathrm{R}},t} := \frac{E_t^{\mathrm{R}}(\sup \mathbb{I}_t)}{E_{t-1}^{\mathrm{R}}(\sup \mathbb{I}_{t-1})} = G_{A^*,t}^{-\zeta^{\mathrm{R}}} G_{\mathbb{I},t}^{\phi^{\mathrm{R}}} = G_{A^*,t}^{-\zeta^{\mathrm{R}}} G_{N,t}^{\phi^{\mathrm{R}}} G_{E^{\mathrm{R}},t}^{\phi^{\mathrm{R}}}$, i.e. $G_{E^{\mathrm{R}},t} = \left(G_{A^*,t}^{-\zeta^{\mathrm{R}}} G_{N,t}^{\phi^{\mathrm{R}}}\right)^{\frac{1}{1-\phi^{\mathrm{R}}}}$.
- **Free entry of inventors:** Either $G_{\mathbb{I},t} \geq 1-\delta_{\mathbb{I},t}$ binds or $\mathcal{L}_t^{\mathrm{I}}\widehat{W}_t^{\mathrm{W}} \geq \widehat{V}_t^{\mathbb{I}}$ does.

7.5.5. Market clearing

- **Labour market clearing:** $\hat{L}_t^{\mathrm{S}} = \mathcal{L}_t^{\mathrm{I}}\hat{\mathbb{I}}_t\left[1-(1-\delta_{\mathbb{I},t})\frac{1}{G_{\mathbb{I},t}}\right] + \eth_t \hat{\mathbb{I}}_t J_t^{\mathrm{P}}\mathcal{L}_t^{\mathrm{RP}} + \widehat{X}_t\left[\left(\frac{1}{\hat{A}_t}\right)^{\frac{1}{\lambda_t}}\left(\frac{1+\mu_{t-1}}{1+\mu_{t-1}^{\mathrm{P}}}\right)^{\frac{1+\lambda}{\lambda}}\eth_{t-1} + \left(\frac{\hat{A}_t^{\mathrm{N}}}{\hat{A}_t}\right)^{\frac{1}{\lambda_t}}\left(\frac{1+\mu_{t-1}}{1+\eta\lambda}\right)^{\frac{1+\lambda}{\lambda}}(1-\eth_{t-1})\right]$.
- **Goods market clearing:** $\widehat{X}_t = \hat{C}_t$.
- **Total output:** $\hat{Y}_t^{\mathrm{W}} = \hat{C}_t + \mathcal{L}_t^{\mathrm{I}}\hat{\mathbb{I}}_t\left[1-(1-\delta_{\mathbb{I},t})\frac{1}{G_{\mathbb{I},t}}\right]\widehat{W}_t^{\mathrm{W}} + \eth_t\hat{\mathbb{I}}_t J_t^{\mathrm{P}}\mathcal{L}_t^{\mathrm{RP}}\widehat{W}_t^{\mathrm{W}}$, where $\hat{Y}_t^{\mathrm{W}} := \frac{Y_t^{\mathrm{W}}}{N_t A_t}$.

[49] As a consequence, we have that $G_{A,t} = \frac{\hat{A}_t}{\hat{A}_{t-1}} G_{A^*,t}$.

[50] This means $G_{I,t} = G_{N,t} G_{A^*,t}^{-\zeta} \frac{\hat{I}_t}{\hat{I}_{t-1}}$.