Aggregate Implications of Innovation Policy

Andrew Atkeson
UCLA

Ariel Burstein
UCLA*

February 23, 2015

Abstract

We examine the quantitative impact of policy-induced changes in innovative investment by firms on growth in aggregate productivity and output in a fairly general specification of a model of growth through firms' innovations that nests several commonly used models in the literature. We present simple analytical results isolating the specific features and/or parameters of the model that play the key roles in shaping its quantitative implications for the aggregate impact of policy-induced changes in innovative investment in the short-, medium- and long-term. We find that the implicit assumption made commonly in models in the literature that there is no *social depreciation of innovation expenditures* plays a key role not previously noted in the literature. Specifically, we find that the elasticity of aggregate productivity and output over the medium-term horizon (i.e. 20 years) with respect to policy-induced changes in the innovation intensity of the economy cannot be large if the model is calibrated to match a moderate initial growth rate of aggregate productivity and builds in the assumption of no social depreciation of innovation expenditures. In this case, the medium-term dynamics implied by the model are largely disconnected from the parameters of the model that determine the model's long run implications and the socially optimal innovation intensity of the economy.

*This is a substantially revised version of a previous draft with the same title, NBER Working Paper 17493. We thank Ufuk Akcigit, Philippe Aghion, Costas Arkolakis, Arnaud Costinot, Ellen McGrattan, Juan Pablo Nicolini, Pedro Teles, Aleh Tsyvinski and Ivan Werning for very useful comments. We thank the National Science Foundation (Award Number 0961992) for research support.

1 Introduction

Firm's investments in innovation are large relative to GDP and are likely an important factor in accounting for economic growth over time.[1] Many OECD countries use taxes and subsidies to encourage these investments in the hope of stimulating economic growth.[2] To what extent can we change the path of macroeconomic growth over the medium and long term if we were to succeed in using innovation policies to induce firms to increase their investments in innovation?

We examine this question in a model of growth through firms' investments in innovation that nests several of the important models of the interaction of firms' investments in innovation and aggregate productivity growth that have been developed over the past 25 years. The models that we nest include the aggregate model of Jones (2002), Neo-Schumpeterian models based on the Quality Ladders framework such as those described in Grossman and Helpman (1991b), Aghion and Howitt (1992), Klette and Kortum (2004) and Lentz and Mortensen (2008), and models based on the Expanding Varieties framework such as those described in Grossman and Helpman (1991a), Luttmer (2007), and Atkeson and Burstein (2010).[3] As described in Aghion et al. (2013), these are influential models that link micro data on firm dynamics to firms' investments in innovation and, in the aggregate, to economic growth in a tractable manner.[4] One of the features that distinguishes these models from standard models of capital accumulation by firms is that there are, potentially, large gaps between the social and the private returns to firms' investments in innovation. Thus, when using these models to study the impact of innovation policy-induced changes in firms' investments in innovation on aggregate growth, one cannot use standard growth accounting methods based on private and social returns and depreciation rates being equated.[5]

[1]There is a wide range of estimates of the scale of firms' investments in innovation. In the new National Income and Product Accounts revised in 2013, private sector investments in intellectual property products were 3.8% of GDP in 2012. Of that amount, Private Research and Development was 1.7% of GDP. The remainder of that expenditure was largely on intellectual property that can be sold such as films and other artistic originals. Corrado et al. (2005) and Corrado et al. (2009) propose a broader measure of firms' investments in innovation, which includes non-scientific R&D, brand equity, firm specific resources, and business investment in computerized information. These broader investments in innovation accounted for roughly 13% of non-farm output in the U.S. in 2005.

[2]See, for example, Chapter 2 of "Economic Policy Reform: Going for Growth", OECD, 2009.

[3]See Acemoglu (2009) for a textbook presentation of many of these models.

[4]Several authors (see, for example, Akcigit and Kerr 2010, Acemoglu et al. 2013and Lentz and Mortensen 2008) have shown that extended versions of the basic Klette-Kortum Neo-Schumpeterian model provide a good fit to many features of micro data on firms.

[5]There is a very large literature that seeks to use standard methods from growth accounting to capitalize firms' investments in innovation and to use the dynamics of that intangible capital aggregate to account for the dynamics of aggregate productivity and output. See, for example, Griliches, ed (1987), Kendrick

In this paper we present simple analytical results approximating the cumulative impulse responses of the logarithm of aggregate productivity and GDP with respect to a policy-induced change in the logarithm of firms' spending on innovation relative to GDP. In the spirit of growth accounting, our approach is to study directly the model technology that links innovative investments by firms to growth in aggregate productivity to isolate the specific features and/or parameters of the model that play the key roles in shaping its quantitative implications.[6] We also use these analytical impulse response functions to highlight the features of the model that drive its implications for the socially optimal innovation intensity of the economy.

We show that under a set of assumptions on the model technology linking firms' innovative investments to growth in aggregate productivity that are satisfied by the most tractable specifications of the models we nest, the aggregate dynamics of aggregate productivity and output induced by permanent policy-induced changes in the innovation intensity of the economy can be summarized by two sufficient statistics: the *impact elasticity of aggregate productivity growth with respect to an increase in aggregate inputs to research,* and *the degree of intertemporal knowledge spillovers in research.* The first of these statistics, the impact elasticity, measures the immediate payoff in terms of productivity growth from an increase in the innovation intensity of the economy. The second of these statistics measures the extent to which a permanent increase in spending on innovation results in a persistent increase in real inputs into innovation. If there are substantial intertemporal knowledge spillovers, then it is feasible for an economy to invest a greater share of output into innovative investment on a permanent basis without running into diminishing returns marked by an increasing price of real research innovations. If these spillovers are weak, then it is not possible to do so — increased aggregate spending on innovative investment simply leads to an increased price of real research innovations in the long run.

We show analytically that one of the key features of our model that drive its quantitative implications for the impact elasticity is the implicit assumption that there is no *social depreciation of innovation expenditures*. We define this social depreciation rate as the counterfactual growth rate of aggregate productivity that would obtain if all firms in the economy invested nothing in innovation.[7] Under our baseline set of assumptions, this

(1994), Griliches (1998), and Corrado and Hulten (2013). Relatedly, McGrattan and Prescott (2012) use an overlapping generations model augmented to include firms' investments in intangible capital to ask how changes in various tax and transfer policies will impact the accumulation of intangible capital and aggregate productivity and GDP.

[6] In this sense our approach is close to Jones (2002).

[7] In Neo-Schumpeterian growth models there is private depreciation of past investments in innovation in terms of their impact on firms' profits — firms gain and lose products and/or profits as they expend resources to innovate. However, an implicit assumption is that there is no social depreciation of these

impact elasticity is bounded by the gap between the baseline growth rate to which the model is calibrated less the social depreciation rate of innovation expenditures. Thus, if one builds in the implicit assumption that there is no social depreciation of innovation expenditures to our model and applies the model to study advanced economies, then our model's quantitative implications for the impact elasticity of aggregate productivity growth are tightly constrained by the low baseline growth rate of aggregate productivity typically observed in these advanced economies. Under the same assumptions, the elasticity of aggregate productivity and output over the medium term horizon (i.e. 20 years) with respect to policy induced changes in the innovation intensity of the economy is not very large and is not very sensitive to changes in the intertemporal knowledge spillovers that determine the long-run implications of the model and the potential welfare gains that might be achieved from a sustained increase in innovation subsidies.[8] We show that, in contrast, if one makes the alternative assumption that past innovations experience even moderate social depreciation, then the model can produce significantly larger medium term elasticities of aggregate productivity with respect to policy-induced changes in the innovation intensity of the economy.

We then show analytically that the welfare implications of the model for the optimal innovation intensity of the economy are determined primarily by its long run implications, which are governed by the degree of intertemporal knowledge spillovers in research and the patience of consumers. If intertemporal knowledge spillovers are large, then a permanent increase in the innovation intensity of the economy can generate very large increases in aggregate productivity and output in the long run. If consumers are patient, they view this long run payoff as having a large benefit relative to the short and medium run cost of lowering current consumption to allow for increased investment in innovation. These analytical results together imply that if there is no social depreciation

investments in terms of their cumulative impact on aggregate productivity — the contribution of past innovation expenditures to aggregate production possibilities never dies-out over time. In contrast, Expanding Varieties models typically assume that there is private and social depreciation of innovation expenditures in the form of product exit. Corrado and Hulten (2013), Aizcorbe et al. (2009) and Li (2012) discuss comprehensive estimates of the depreciation rates of innovation expenditures without distinguishing between measures of the private and social depreciation of these expenditures.

[8] Comin and Gertler (2006) develop a model of medium-term business cycles based on endogenous movements in aggregate productivity that includes adoption, variable markups, and variable factor utilization. They find that with the combination of these factors, their model can account for significant medium-term cyclical productivity dynamics. We see our results as highlighting the endogenous dynamics of aggregate productivity that arise solely from policy-induced variation in the innovation intensity of the economy. McGrattan and Prescott (2012) emphasize how measurement conventions for GDP impact the measurement of aggregate productivity in the face of time variation in the scale of firms' investments in intangible capital. We discuss the role of different NIPA convention methods for shaping the responses of measured productivity and output.

of knowledge and the initial baseline growth rate of aggregate productivity is low, then the medium term dynamics implied by the model are largely disconnected from the parameters of the model that determine the model's long run implications and the socially optimal innovation intensity of the economy.

Our model allows for a rich and yet tractable model of the birth, growth, and death of firms in which these firm dynamics are driven by incumbent and entrant firms' investments in innovation. For example, one of the principal innovations of the Klette and Kortum (2004) model relative to the standard Quality Ladders model introduced by Aghion and Howitt (1992) and Grossman and Helpman (1991b) is that it considers investments in innovation by both incumbent and entering firms. We show that the two models (both of which are nested in our model), if calibrated equivalently imply the same response of aggregate productivity and output to a change in the innovation intensity of the economy in the long run, up to a first-order approximation. The transition, however, is faster in the Quality Ladders model than it is in the Klette Kortum model. We show more generally that this result follows from the assumption in the Klette-Kortum model that incumbent firms have a smaller average cost of innovation than entering firms.

To focus attention on the aggregate implications of innovation policies that change the aggregate innovation intensity of the economy, we maintain a baseline set of assumptions which imply that, while the aggregate level of innovation expenditures may be suboptimal, there is no misallocation of innovation expenditures across firms in the model economy at the start of the transition following a change in innovation policies. Thus we abstract from the role innovation policies might play in improving the allocation of innovation expenditures across firms and thus raising the aggregate innovation rate without increasing aggregate innovation expenditures. There is a growing literature examining this possibility, see for example Acemoglu et al. (2013), Buera and Fattal-Jaef (2014), Lentz and Mortensen (2014), and Peters (2013). We see our results as providing a useful analytical benchmark to which numerical results from richer models can be compared. We illustrate such a comparison using the recent parameterization of Lentz and Mortensen (2014) in a rich estimated model.

The paper is organized as follows. Section 2 describes the model, the key assumptions of our analytic results, how four specific model examples are nested by our framework. Section 3 characterizes a balanced growth path. Section 4 presents analytic results on the impact of changes in innovation policy on aggregate outcomes at different horizons. Section 5 discusses the quantitative implications of our analytic results. Section 6 concludes. The appendix provides some proofs and other details including the calibration and full numerical solution of the model.

2 Model

In this section we first describe the physical environment that allows for a fairly technology mapping the innovation choices of individual firms to the growth of aggregate productivity. We then describe four specific model examples that fit into our framework. Finally we present aggregate equilibrium conditions that we use when deriving our analytic results.

2.1 Physical environment

Time is discrete and labeled $t = 0, 1, 2, ...$ There are two final goods, the first which we call the *consumption good* and the second which we call the *research good.* The representative household has time-additive preferences over consumption per capita C_t/L_t given by $\sum_{t=0}^{\infty} \frac{\beta^t}{1-\xi} L_t (C_t/L_t)^{1-\xi}$, with $\beta \leq 1$, $\xi > 0$, and where L_t denotes the population that (without loss of generality for our results) is constant and normalized to 1 ($L_t = 1$). The consumption good is produced as a constant elasticity of substitution (CES) aggregate of the output of a continuum of differentiated intermediate goods. These intermediate goods are produced by firms using capital and labor. Labor can be allocated to current production of intermediate goods, L_{pt}, and to research, L_{rt}, subject to the resource constraint $L_{pt} + L_{rt} = L_t$.

Output of the consumption good, Y_t, is used for two purposes. First, as consumption by the representative household, C_t. Second, as investment in physical (tangible) capital, $K_{t+1} - (1 - d_k) K_t$, where K_t denotes the aggregate physical capital stock and d_k denotes the depreciation rate of physical capital. The resource constraint of the final consumption good is

$$C_t + K_{t+1} - (1 - d_k) K_t = Y_t. \tag{1}$$

Under the national income and product accounting (NIPA) convention that expenditures on innovation are expensed, the quantity in the model corresponding to Gross Domestic Product (GDP) as measured in the data under historical measurement procedures is equal to Y_t.[9]

[9] The treatment of expenditures on innovation in the NIPA in the United States is being revised as of the second half of 2013 to include a portion of those expenditures on innovation in measured GDP. If all intangible investments in the model were measured as part of GDP, then measured GDP would be given by $GDP_t = C_t + K_{t+1} - (1 - d_k) K_t + P_{rt} Y_{rt}$, where $P_{rt} Y_{rt}$ denotes intangible investment expenditure, as defined below. We report results on GDP under both measurement procedures.

Intermediate Goods Producing Firms: Intermediate goods (which are used to produce the final consumption good) are produced by heterogeneous firms. Production of an intermediate good with productivity index z is carried out with physical capital, k, and labor, l, according to

$$y = \exp(z) k^{\alpha} l^{1-\alpha}, \tag{2}$$

where $0 < \alpha < 1$.

To maintain a consistent notation across a potentially broad class of models, we assume that there is a countable number of types of firms indexed by $j = 0, 1, 2, \ldots$. As a matter of convention, let $j = 0$ indicate the type of entering firms (we assume that there is only one type of entering firm), and let $j = 1, 2, \ldots$ indicate the potentially different types of incumbent firms. The type of a firm, j, records all the information about the firm regarding the number of intermediate goods it produces, the different productivity indices z with which it can produce these various goods, the markups it charges for each of these goods that it produces, and all the relevant information about the technologies the firm has available to it for innovating. Specifically, focusing on the production technology, firm $j \geq 1$ is the owner of the frontier technology for producing $n(j)$ intermediate goods with vector of productivities $(z_1(j), z_2(j), \ldots, z_{n(j)}(j))$. The type of the firm j also records the equilibrium markups of price over marginal cost $(\mu_1(j), \mu_2(j), \ldots, \mu_{n(j)}(j))$ that this firm charges on the goods that it produces. We illustrate below how the firm type j also records the technology for innovation available to the firm. We provide below four examples of how existing models fit into this notational framework.

Let $\{N_t(j)\}_{j \geq 1}$ denote the measure of each type j of incumbent firm at time t. The measure of intermediate goods being produced in the economy at date t is $\sum_{j \geq 1} n(j) N_t(j)$. The vector of types of incumbent firms $\{N_t(j)\}_{j \geq 1}$ is a state variable that evolves over time depending on entry and the investments in innovative activity of the incumbent firms, as described below.

Production of the final consumption good: Letting $y_{kt}(j)$ denote the output of the k'th product of firm type j at time t, then output of the final good is given by

$$Y_t = \left(\sum_{j \geq 1} \sum_{k=1}^{n(j)} y_{kt}(j)^{(\rho-1)/\rho} N_t(j) \right)^{\rho/(\rho-1)}. \tag{3}$$

with $\rho \geq 1$. This technology for producing the consumption final good is operated by competitive firms, with standard demand functions for each of the intermediate goods.

We assume that within each period, capital and labor are freely mobile across products

and intermediate goods producing firms. This implies that the marginal cost of producing the k'th product of firm type j at time t is given by $MC_t \exp(-z_k(j))$, where MC_t is the standard unit cost for the Cobb-Douglass production function (2) with $z = 0$. We assume that this firm charges price $p_{kt}(j) = \mu_{kt}(j) MC_t \exp(-z_k(j))$ for this product at markup $\mu_{kt}(j)$ over marginal cost. This gives us that aggregate output can be written as

$$Y_t = Z_t (K_t)^{\alpha} (L_{pt})^{1-\alpha}, \tag{4}$$

where L_{pt} and K_t are the aggregates across intermediate goods producing firms of labor and capital used in current production and Z_t corresponds to *aggregate productivity* in the production of the final consumption good:[10]

$$Z_t \equiv Z(\{N_t(j)\}_{j\geq 1}) = \frac{\left(\sum_{j\geq 1}\sum_{k=1}^{n(j)} \exp((\rho-1)z_k(j))\mu_k(j)^{1-\rho} N_t(j)\right)^{\rho/(\rho-1)}}{\sum_{j\geq 1}\sum_{k=1}^{n(j)} \exp((\rho-1)z_k(j))\mu_k(j)^{-\rho} N_t(j)} \tag{5}$$

When markups are equal across firms, this expression for aggregate productivity simplifies to

$$Z_t \equiv Z(\{N_t(j)\}_{j\geq 1}) = \left(\sum_{j\geq 1}\sum_{k=1}^{n(j)} \exp((\rho-1)z_k(j)) N_t(j)\right)^{1/(\rho-1)}. \tag{6}$$

The research good: Intermediate goods' producing firms use the second final good, which we call the research good and whose production is described below, to invest in innovative activities. Let $\{y_{rt}(j)\}_{j\geq 1}$ denote the use of the research good by each type j of incumbent firms in period t. We can model the use of the research good by entering firms in two possible ways. We can consider the mass of entering firms as fixed over time as a parameter, $\bar{N}(0)$, and allow the use of the research good by these entering firms $y_{rt}(0)$ to vary. Alternatively, we can consider the use of the research good by each entering firm to be fixed as a parameter at $\bar{y}_r(0)$ and let the mass of entering firms $N_t(0)$ to vary. For simplicity we use the second convention and discuss explicitly where in our results this choice makes a difference. Given this notation, the resource constraint for the research good in each period t is given by

[10] In general, this model-based measure of aggregate productivity, Z_t, does not correspond to measured TFP, which is given by $TFP_t = GDP_t / \left(K_t^{\tilde{\alpha}} L_t^{1-\tilde{\alpha}}\right)$, where $1-\tilde{\alpha}$ denotes the share of labor compensation in measured GDP. This adjustment is required because of the expensing of expenditures on innovation (under historical standards for measuring GDP) and because of possible variation over time in the allocation of labor between production and research. The growth rate of this model-based measure of aggregate productivity, however, is equal to the growth rate of measured TFP on a balanced growth path (in terms of aggregates).

$$\sum_{j\geq 1} y_{rt}(j)N_t(j) + \bar{y}_r(0)N_t(0) = Y_{rt}, \tag{7}$$

where Y_{rt} denotes the aggregate output of the research good.

Production of the research good is carried out using research labor L_{rt} according to

$$Y_{rt} = Z_t^{\gamma-1} A_{rt} L_{rt} \ , \tag{8}$$

where $\gamma \leq 1$. The variable A_{rt} represents the stock of basic scientific knowledge that is assumed to evolve exogenously, growing at a steady rate of $g_{A_r} \geq 0$ so $A_{rt+1} = \exp(g_{A_r})A_{rt}$. Increases in this stock of scientific knowledge improve the productivity of resources devoted to innovative activity. We interpret A_{rt} as a worldwide stock of scientific knowledge that is freely available for firms to use in innovative activities. The determination of this stock of scientific knowledge is outside the scope of our analysis.[11] In the Appendix we consider an extension in which research production uses both labor and consumption good, as in the lab-equipment model of Rivera-Batiz and Romer (1991).

We interpret the parameter $\gamma \leq 1$ as indexing the extent of *intertemporal knowledge spillovers,* that is the extent to which further innovations become more difficult as the aggregate productivity Z_t grows relative to the stock of scientific knowledge A_{rt}. As γ approaches 1, the resource cost of innovating on the frontier technology becomes independent of Z_t. The impact of advances in Z_t on the cost of further innovations is external to any particular firm and hence we call it a spillover. Standard specifications of quality ladders models with fully endogenous growth correspond to the case with full spillovers and $g_{A_r} = 0$.

Innovation by firms: Aggregate productivity in the model grows as a result of the investment in innovation by firms. We model the technology for innovation abstractly as a transition law for the distribution of incumbent firms as follows. Given a collection of incumbent firms $\{N_{jt}\}_{j\geq 1}$ and investments in innovation by these firms $\{y_{rt}(j)\}_{j\geq 1}$ as well as a measure of entering firms N_{0t} in period t, the types of all firms are updated giving a

[11] It is common in the theoretical literature to assume that all productivity growth is driven entirely by firms' expenditures on R&D (Griliches 1979, p. 93). As noted in Corrado et al. (2011), this view ignores the productivity-enhancing effects of public infrastructure, the climate for business formation, and the fact that private R&D is not all there is to innovation. We capture all of these other productivity enhancing effects with A_r. Relatedly, Akcigit et al. (2013) considers a growth model that distinguishes between basic and applied research and introduces a public research sector.

new collection of incumbent firms at $t+1$, $\{N_{t+1}(j)\}_{j\geq 1}$, given by

$$\{N_{t+1}(j)\}_{j\geq 1} = T\left(\{y_{rt}(j)\}_{j\geq 1}, N_t(0); \{N_t(j)\}_{j\geq 1}\right). \tag{9}$$

By equations (5) and (9), the models we consider deliver an equation that gives the growth rate of aggregate productivity as a function of entry and the use of the research good by all types of incumbent firms,

$$g_{zt} \equiv \log(Z_{t+1}) - \log(Z_t) = G\left(\{y_{rt}(j)\}_{j\geq 1}, N_t(0); \{N_t(j)\}_{j\geq 1}\right), \tag{10}$$

where g_{zt} is the symbol we use to denote the growth of log aggregate productivity between t and $t+1$. Throughout we assume that the function G is differentiable and that its domain is a convex set. We denote by $G_t^0 = G(\{0,0,\ldots\}, 0; \{\bar{N}_t(j)\}_{j\geq 1})$ the growth rate of aggregate productivity if all firms, both entrants and incumbents, were to set their use of the research good to zero. We refer to G_t^0 as the *social depreciation rate of innovation* expenditures.[12]

Policies: Intermediate goods producing firms are offered type-specific innovation subsidies $\tau(j)$. A firm of type j that purchases $y_{rt}(j)$ units of the research good at time t pays $P_{rt}y_{rt}(j)$ to a research good producer for that purchase and then receives a rebate of $\tau_t(j) P_{rt}y_{rt}(j)$ from the government. Thus fiscal expenditures on these policies are given by $E_t = \sum_{j\geq 1} \tau_t(j) y_{rt}(j) N_t(j) + \tau_t(0) \bar{y}_r(0) N_t(0)$. Changes in innovation policies are then assumed to lead to changes in the equilibrium allocation of the research good across firms and hence aggregate productivity growth and the time path for all other macroeconomic variables. These changes in innovation policies are assumed not to directly effect the functions Z, T, and G defined above in equations (5), (9), and (10).

Example models and the corresponding G function

The function G describing the law of motion of aggregate productivity as a function of firms' investments in innovation plays a central role in our analysis. We find it useful to make reference to the following four example models and to derive the function G in each case.[13]

[12] This is analogous to defining the depreciation of physical capital as $\log(1-d_k)$. In defining G_t^0, we are assuming that $y_{rt}(j) = N_t(0) = 0$ for all $j \geq 1$ is in the domain of G.

[13] While we do not present their model as one of our four specific examples, the model of Bloom et al. (2013) with innovation by firms with technological spillovers to neighboring firms is nested in our framework.

Example 1: Simple quality ladders model In the quality ladders model of Grossman and Helpman (1991b) and Aghion and Howitt (1992), incumbent firms each produce one good at a constant markup and are thus indexed by the productivity with which they can produce this good, $\exp(z)$. The index of firm productivity has countable support (the ladder) that we will index as $z(j)$ for $j = 1, 2, \ldots$. The measure of entering firms at each date, $N_t(0)$ is an endogenous variable. With a fixed measure of intermediate goods of size 1, there is a measure 1 of incumbents, $\sum_{j\geq 1} N_t(j) = 1$ for all t. Incumbents do not expend resources on innovation, so $y_{rt}(j) = 0$ for all t and $j \geq 1$. Under these assumptions, with constant markups across products, aggregate productivity is given by (6).

The innovation technology in this model is as follows. Each entering firm at t produces in expectation σ innovations. Each innovation is matched randomly (uniformly) to an existing intermediate good (each intermediate good receives at most one innovation every period). This innovation raises the productivity with which that good can be produced from $\exp(z(j))$ to $\exp(z(j+1)) = \exp(z(j) + \Delta_z)$. With such a successful innovation on a product produce by a firm of type j, the firm entering at t becomes a firm of type $j+1$ at $t+1$ and the previous incumbent firm of type j at t ceases operating. The total measure of products innovated on is $\sigma N_t(0)$, which is assumed to be less than one (to ensure that this condition is satisfied in a discrete time model, one can simply reduce the length of a time period). Given the definition of aggregate productivity in equation (6), we can write the function G in equation (10) simply as a function of the measure of entering firms $N_t(0)$ and parameters:

$$G\left(\{y_{rt}(j)\}_{j\geq 1}, N_t(0); \{N_t(j)\}_{j\geq 1}\right) = \frac{1}{\rho - 1} \log\left(\sigma N_t(0)\left(\exp(\Delta_z)^{\rho-1} - 1\right) + 1\right) \quad (11)$$

In this model, there is no social depreciation of innovation expenditures, i.e. $G_t^0 = 0$.

Example 2: Simple expanding varieties model Consider now an expanding variety model similar to that in Luttmer (2007). As in the Quality Ladders model, we assume that incumbents produce a single product and are indexed by the productivity with which they can produce it, $\exp(z)$. Markups are again constant across firms and over time. Hence, aggregate productivity is given by equation (6).

The productivity of incumbents grows exogenously for those incumbents that survive. Specifically, each incumbent firm has exogenous probability δ_f of exiting the market each period. If an incumbent of type $j \geq 1$ at t, with productivity $\exp(z(j))$, survives to period $t+1$, then it becomes type $j+1$ at that date and has productivity $\exp(z(j+1)) = \exp(z(j) + \Delta_z)$. Incumbents are assumed not to expend resources on innovation,

so $y_{rt}(j) = 0$ for all t and $j \geq 1$. In the Appendix we show how to extend this example to allow for innovation by incumbents, as in e.g. Atkeson and Burstein (2010), along the lines of Example 3.

We no longer assume that the measure of goods is fixed at one as in the Quality Ladders model. Instead, the measure of goods in production can expand or contract over time due to exogenous exit of incumbents (δ_f) and the entry of new firms. Hence aggregate productivity can grow both due to the exogenous productivity growth in surviving incumbent firms and due to endogenous entry of new firms producing new products. There is a spillover of knowledge from incumbents to entrants, such that entrants at time t start production at time $t+1$ with productivity drawn from a lottery with probabilities $\eta(j; Z_t)$ over types (values of $z(j)$) that satisfies$\sum_j \exp(z(j))^{(\rho-1)} \eta(j; Z_t) = \lambda Z_t^{\rho-1}$.[14] In this case we can write the function G in equation (10) as

$$G\left(\{y_{rt}(j)\}_{j\geq 1}, N_t(0); \{N_t(j)\}_{j\geq 1}\right) = \frac{1}{\rho-1}\log\left(\left(1-\delta_f\right)\exp\left(\left(\rho-1\right)\Delta_z\right) + \lambda N_t(0)\right).$$

In this model, as long as incumbent firms shrink as a group in the sense that $\left(1-\delta_f\right)\exp\left(\left(\rho-1\right)\Delta_z\right) < 1$ there is social depreciation of innovation expenditures, i.e. $G^0 = \frac{1}{\rho-1}\log\left(\left(1-\delta_f\right)\exp\left(\left(\rho-1\right)\Delta_z\right)\right) < 0$. The magnitude of the social rate of depreciation of innovation expenditures in this model is directly linked to the model's calibrated value for the employment share of incumbent firms, which is given by $\exp((\rho-1)G^0)/\exp((\rho-1)\bar{g}_z)$.

Example 3: Simple Quality ladders model with innovation by incumbents (based on Klette and Kortum 2004): This model is an extension of the simple Quality Ladders model in which both entrant and incumbent firms expend the research good in an effort to innovate. Incumbent firms are indexed now by a vector $j = (n(j), z_1(j), z_2(j), \ldots, z_{n(j)}(j))$, for $j \geq 1$ indicating the number of products $n(j)$ for which this firm is the frontier producer and the vector of productivities with which this firm can produce these products. Note that since we have a continuum of measure one of products, then $\sum_{j\geq 1} n(j)N_t(j) = 1$ for all t for all feasible allocations. Under the assumption that all firms choose a constant markup on all products that they produce and that this markup is constant over time, aggregate productivity is again given by equation (6).

[14] Luttmer (2007) considers an alternative specification of the distribution of productivities of entrants, $\eta(j; Z_t) = N_t(j)/\sum_{j\geq 1} N_t(j)$. With this specification of the knowledge spillover from incumbents to entrants, it may be possible to raise the contribution of entry to the growth rate of aggregate productivity by eliminating incumbent firms with below average productivity. As shown by Acemoglu et al. (2013), a policy induced reduction in the number of incumbent firms may be welfare improving despite the cost of decreasing the current value of aggregate productivity Z_t.

Incumbent firms of type j that own $n(j)$ products have an innovation technology such that if they expend $y_{rt}(j)$ units of the research good, the have $\sigma d(y_{rt}(j)/n(j))n(j)$ innovations in expectation, where $d(.)$ is an increasing and concave function. Each innovation is matched randomly (uniformly) to an existing intermediate good (owned by some other firm) raising its frontier productivity from $\exp(z(j))$ to $\exp(z(j+1)) = \exp(z(j)+\Delta_z)$. We assume that $y_{rt}(j)/n(j) = y_{rt}(1)$ for all $j \geq 1$ (where $j = 1$ indexes a firm with one product and $z_1(1) = 1$, the lowest rung on the quality ladder). In the appendix we show that this is an equilibrium outcome if all incumbents face the same proportional innovation subsidies (not necessarily equal to the subsidy faced by entrants) and charge the same markups. The total measure of products innovated on by incumbents at time t under this assumption is $\sigma d(y_{rt}(1)) \sum_{j\geq 1} n(j)N_t(j) = \sigma d(y_{rt}(1))$.

As in the standard quality ladders model, we have that entrants expend resource $\bar{y}_r(0)$ to have σ innovations in expectations, so that the total measure of product innovated on by entrants is $\sigma N_t(0)$.

Under these assumptions, we have that the function G can be written

$$G\left(\{y_{rt}(j)\}_{j\geq 1}, N_t(0); \{N_t(j)\}_{j\geq 1}\right) = \frac{1}{\rho-1}\log\left(\sigma(d(y_{rt}(1)) + N_t(0))\left(\exp(\Delta_z)^{\rho-1} - 1\right) + 1\right) \tag{12}$$

As in the basic quality ladders model, in this model there is no social depreciation of innovation expenditures, i.e. $G_t^0 = 0$. In the appendix we present additional details of this model example.

Example 4: Klette Kortum 2004 model with innovation by incumbents and heterogeneous step sizes (Lentz and Mortensen 2014): The Lentz and Mortensen (2008) and Lentz and Mortensen (2014) model is a quantitative implementation of the Klette-Kortum model in which incumbent firms all face the same cost function for producing innovations $d(y_{rt}(j))$, but high quality innovation type firms produce innovations with a larger increment to the productivity of the product on which the innovation occurs (Δ_{zH}) while low quality innovation type firms produce innovations with a smaller increment to the productivity of the product on which the innovation occurs (Δ_{zL}).

Incumbent firms are indexed now by a vector $j = (m(j), n(j), z_1(j), z_2(j), \ldots, z_{n(j)}(j))$, for $j \geq 1$ indicating the innovation quality type $m(j) \in \{H, L\}$, the number of products $n(j)$ for which this firm is the frontier producer and the vector of productivities with which this firm can produce these products. Note that since we have a continuum of measure one of products, $\sum_{j\geq 1} n(j)N_{jt} = 1$ for all t for all feasible allocations. We let N_t^H

denote the measure of products produced by high quality innovation firms at time t, that is

$$N_t^H = \sum_{j|m(j)=H} n(j)N_t(j)$$

and likewise $N_t^L = 1 - N_t^H$. We assume that entering firms draw an innovation quality type immediately with their first successful innovation. These entering firms have probability ϕ^H of drawing the high quality type and ϕ^L the low quality type. Note that in the previous examples we have assumed that each entering firm requires a fixed expenditure of resources $\bar{y}_{rt}(0)$. Lentz and Mortensen (2014) instead assume that the measure of entering firms $N_t(0)$ is fixed at $\bar{N}(0)$ and that each entrant has access to the same technology for generating innovations as incumbents. Thus, each entrant has an endogenous probability $\sigma d(y_{rt}(0))$ of producing an innovation. This leads to a slight alteration of our notation.

Assuming that all incumbents firms within an innovation type face the same innovation policies and charge the same markup, equilibrium allocations are such that for all incumbent firms of the high quality innovation type (for all $j \geq 1$ with $m(j) = H$), $y_{rt}(j)/n(j)$ is equal to $y_{rt}^H(1)$ and likewise $y_{rt}^L(1)$ for low innovation quality firms. Entrants all choose to expend resources $y_{rt}(0)$ in period t.

For simplicity, assume that the markups that firms charge in the product market is independent of their innovation type.[15] Under this assumption, we again have aggregate productivity given as in equation (6). Under these assumptions, the growth rate of aggregate productivity is given by

$$G\left(\{y_{rt}(j)\}_{j\geq 0}; \{N_t(j)\}_{j\geq 0}\right) = \tag{13}$$

$$\frac{1}{\rho-1}\log\left(\sum_{m=H,L}\sigma(d(y_{rt}^m(1))N_t^m + d(y_{rt}(0))\phi^m\bar{N}(0))\left(\exp(\Delta_z^m)^{\rho-1}-1\right)+1\right).$$

As in the basic quality ladders model, in this model there is no social depreciation of innovation expenditures, i.e. $G_t^0 = 0$.

[15] This will be the case if the innovation step size for both types of firms exceeds the monopoly markup or if the latent producer on each product has a technology a fixed step size Δ_l behind each incumbent producer. If we do not make this assumption of constant markups, then changes in the distribution of markups due to changes in the distribution of firm types will have an impact on aggregate productivity growth in the transition of this economy from one BGP to another and will impact the level of aggregate productivity on a given BGP. We thus abstract here from the impact of innovation policies on aggregate productivity through their impact on the distribution of markups in our examples, as studied by e.g. Peters (2013).

Macroeconomic equilibrium conditions

We assume that the representative household owns the incumbent firms and the physical capital stock, facing a sequence of budget constraints given by

$$C_t + K_{t+1} = \left[R_{kt} + (1 - d_k)\right] K_t + W_t L_t + D_t - E_t,$$

in each period t, where W_t, R_{kt}, D_t, and E_t denote the economy-wide wage (assuming that labor is freely mobile across production of intermediate goods and the research good), rental rate of physical capital, aggregate dividends paid by firms, and aggregate fiscal expenditures on policies (which are financed by lump-sum taxes collected from the representative household), respectively.

Production of the research good is undertaken by competitive firms that do not internalize the intertemporal knowledge spillover from innovation as given. Cost minimization in the production of the research good implies that the price of the research good, P_{rt}, is equal to

$$P_{rt} = \frac{Z_t^{1-\gamma}}{A_{rt}} W_t \,. \tag{14}$$

P_{rt} is the deflator needed to translate changes in innovation expenditure, $P_{rt}Y_{rt}$, into changes in innovation output, Y_{rt}. We define the *innovation intensity of the economy*, s_{rt}, as the ratio of innovation expenditure to the sum of expenditure on consumption and physical capital investment (i.e. the old measure of GDP), that is $s_{rt} = P_{rt}Y_{rt}/GDP_t$. It is typically a challenge to measure real research output Y_{rt}. Instead, the data that are usually available are data on research spending. To compute how production of the research good Y_r changes with changes in expenditure on innovation relative to GDP s_r, we make use of the following results about the division of GDP into payments to various factors of production and the relationship of those factor shares to the innovation intensity of the economy and the allocation of labor.

With CES aggregators and Cobb-Douglas production functions, aggregate revenues of intermediate goods firms, Y_t, are split into three components. A share $\frac{\mu_t - 1}{\mu_t}$ accrues to variable profits from intermediate goods production equal to their total sales (which sum to output of the final consumption good Y_t) less aggregate wages paid to production labor and aggregate rental payments to physical capital. We define μ_t directly this way as a share of aggregate output of the final consumption good and refer to it as the average markup. Of the remaining revenues, a share α/μ_t is paid to physical capital, $R_{kt}K_t = \frac{\alpha}{\mu_t}Y_t$, and a share $(1-\alpha)/\mu_t$ is paid as wages to production labor, $W_t L_{pt} = \frac{(1-\alpha)}{\mu_t}Y_t$.

With perfect competition in the research sector, $W_t L_{rt} = P_{rt}Y_{rt}$. Using the factor shares

above and the assumption that labor is freely mobile between production and research, the allocation of labor between production and research is related to expenditures on the research good by[16]

$$\frac{L_{pt}}{L_{rt}} = \frac{(1-\alpha)}{\mu_t}\frac{1}{s_{rt}}. \tag{15}$$

3 Balanced growth path

We consider balanced growth paths (BGP's) of the following form: the innovation intensity of the economy s_{rt}, the allocation of labor between production and research, L_{pt} and L_{rt}, output of the research good Y_{rt} and the growth rate of aggregate productivity g_{zt} all remain constant over time at the levels $\bar{L}_p$, $\bar{L}_r$, $\bar{Y}_r$ and $\bar{s}_r$, respectively. Whether such a BGP exists or not depends on model details that we have not yet specified. In deriving our analytic results, we assume that such a BGP exists. We then verify this conjecture for the specific model examples we consider.

If such a BGP exists and if $\gamma < 1$, then our model is a semi-endogenous growth model with the growth rate along the BGP determined by the exogenous growth rate of scientific knowledge g_{A_r} and other parameter values independently of innovation policies, as in Kortum (1997) and Jones (2002). In this case, it is not possible to have fully endogenous growth because such growth would require growth in innovation expenditure in excess of the growth rate of GDP. Ongoing balanced growth can occur only to the extent that exogenous scientific progress reduces the cost of further innovation as aggregate productivity Z grows. These BGP growth rates are given from equations (1), (4), and (8) as follows. The growth rate of aggregate productivity is given by $\bar{g}_z = g_{Ar}/(1-\gamma)$. The growth rate of output of the consumption good (and hence consumption, physical capital, and the wage) is given by $\bar{g}_y = \bar{g}_z/(1-\alpha)$, and the rental rate of capital is constant and given by $\bar{R}_k = \beta^{-1}\exp\left(\tilde{\xi}\bar{g}_y\right) - 1 + d_k$.

If a BGP exists and the knife edged conditions $\gamma = 1$ and $g_{Ar} = 0$ hold, then our model is an endogenous growth model with the growth rate along the BGP determined by firms' investments in innovative activity, as in Grossman and Helpman (1991b) and Klette and

[16]Here we are assuming that there is one wage for labor in both production and research. In the Appendix we present an extension in which labor is imperfectly substitutable between production and research as in Jaimovich and Rebelo (2012). The assumption of imperfect substitutability reduces the elasticity of the allocation of labor between production and research with respect to a policy-induced change in the innovation intensity of the economy, resulting in even smaller responses of aggregate productivity and GDP to a given change in the innovation intensity of the economy relative to those in our baseline model. The is similar to assuming congestion in the production of the research good (i.e. in which case research labor in the production of the research good has an exponent less than one), as discussed in Jones (2005).

Kortum (2004).[17] The transition paths of the response of aggregates to policy changes are continuous as γ approaches one. [18]

We will calibrate the model parameters to match a given BGP per capita growth rate of output, $\bar{g}_y$, rather than making assumptions about the growth rate of scientific knowledge, g_{A_r}, which is hard to measure in practice. Specifically, given a choice of $\bar{g}_y$ and physical capital share of α, the growth rate of aggregate productivity in the BGP is $\bar{g}_z = \bar{g}_y (1-\alpha)$. For a given choice of γ, we choose the growth rate of scientific knowledge consistent with this productivity growth rate, that is $g_{Ar} = (1-\gamma)\,\bar{g}_z$.

4 Aggregate implications of changes in the innovation intensity of the economy: analytic results

In this section, we derive analytic results regarding the impact of policy-driven changes in the innovation intensity of the economy on aggregate outcomes at different time horizons. These analytical results demonstrate what features of our baseline model are key in determining its implications for the aggregate impact of innovation policies. In the next section, we discuss the quantitative implications of our analytical results.

In framing the question of how policy-induced changes in the innovation intensity of the economy impact aggregate outcomes at different time horizons, we consider the following thought experiment. Consider an economy that is initially on a BGP with growth rate of aggregate productivity $\bar{g}_z$. As a baseline policy experiment, consider a change in innovation policies to new innovation subsidies beginning in period $t = 0$ and continuing on for all $t > 0$. This policy experiment leads to some observed change in the path of the innovation intensity of the economy $\{s'_{rt}\}_{t=0}^{\infty}$ different from the innovation intensity of the economy $\bar{s}_r$ on the original BGP as well as some reallocation of innovation expenditure across firms and some evolution of the distribution of incumbent firms across firm types. We seek to analytically approximate the quantitative predictions of our model for the associated equilibrium change in the path for model productivity $\left\{Z'_t\right\}_t^{\infty}$ and other macroeconomic aggregates to determine what features of the model are important in shaping the model's quantitative predictions for the aggregate impact of innovation policies on these macroeconomic aggregates.

[17] If $\gamma > 1$, then our model does not have a BGP, as in this case, a constant innovation intensity of the economy leads to an acceleration of the innovation rate as aggregate productivity Z grows.

[18] This intertemporal knowledge spillover parameter γ hence plays the same role as the knowledge spillover parameter ϕ discussed in Section 5 of Jones (2005). He makes the same argument that the specific choice of $\gamma = 1$ or $\gamma < 1$ but close to one does not significantly impact the model's medium term transition dynamics because of continuity of the transition paths in this parameter.

We first approximate and bound quantitatively the elasticity at the start of the transition of the growth rate of aggregate productivity with respect to a change in aggregate use of the research good Y_{rt}. We then approximate the transition dynamics of aggregate productivity and GDP from an initial baseline BGP to the new BGP corresponding to the new set of innovation policies. In deriving these results, we make use of three key assumptions regarding the model implied function G linking use of the research good by entering and incumbent firms to aggregate productivity growth. These assumptions are that: (i) the allocation of research expenditure across firms on the baseline BGP is conditionally efficient as described below, (ii) the function G satisfies at least one of two concavity assumptions described below, and (iii) the state $\{N_t(j)\}_{j\geq 1}$ does not enter the function G. We illustrate the applicability of these assumptions to our model examples 1-3. The most striking aspect of these results is that we are able to develop an upper bound on the quantitative implications of our model for these productivity dynamics using a small set of parameters and sufficient statistics from the model. We then use our analytical approximation of the transition dynamics of productivity and GDP following a permanent change in the innovation intensity of the economy to provide an analytical characterization of the optimal innovation intensity of the economy as a function of the same small set of parameters and the discount rate of consumers that applies for economies that satisfy the same three assumptions.

We use these analytical results in the next section to show that for models that satisfy these assumptions and that assume no social depreciation of innovation expenditures as defined below (e.g. examples 1 and 3), this bound tightly restricts the quantitative implications of the model for the response of aggregate productivity growth to changes in the innovation intensity of the economy in transitions following a change in innovation policies over the short and medium term horizons. We show that for models that assume even a moderate social depreciation rate of innovation expenditures (e.g. example 2), this bound is substantially relaxed.

The analytical results presented in this section abstract from the transition dynamics and welfare gains that might be induced by a reallocation of innovation expenditures across firms when assumptions (i) and/or (iii) are violated, i.e. when the initial allocation is not conditionally efficient and/or when the distribution of incumbent firm types enters into the function G. We discuss how it is possible to violate these assumptions in model examples 3 and 4 and we discuss the additional terms that must be computed to provide a characterization of the quantitative implications of these models in this case. In the next section, we examine the quantitative importance of these additional terms in the estimated version of the model in Example 4 provided in Lentz and Mortensen (2014).

Three key assumptions

We first spell out three key assumptions used in our analytical results.

Assumption 1: Conditional Efficiency of $\{\bar{y}_{rt}(j)\}_{j\geq 1}$ and $\bar{N}_t(0)$: We say that the allocation $\{\bar{y}_{rt}(j)\}_{j\geq 1}$ and $\bar{N}_t(0)$ is *conditionally efficient* if it is an interior solution to the problem of maximizing G subject to the resource constraint (7) when the state variables $\{\bar{N}_t(j)\}_{j\geq 1}$ and total production of the research good $\bar{Y}_{rt}$ are taken as given.

Assumption 2: A concavity assumption on G: We impose that the function G be concave in a certain direction as follows. Let $\bar{Y}_r$, $\{\bar{y}_{rt}(j)\}_{j\geq 1}$, $\bar{N}_t(0)$ with the state variables $\{\bar{N}_t(j)\}_{j\geq 1}$ be the baseline BGP endogenous variables at time t. Assume that the allocation with $y_{rt}(j) = 0$ for all $j \geq 1$ and $N(0) = 0$ is also in the domain of the function G. Define the function $H_t(a)$ over the domain $a \in [0, 1+\epsilon)$ for a fixed $\epsilon > 0$ as equal to G evaluated at the point

$$(1-a) * 0 + a * (\{\bar{y}_{rt}(j)\}_{j\geq 1}, \bar{N}_t(0)),$$

with the state variables $\{\bar{N}_{jt}\}_{j\geq 1}$ taken as fixed. We assume that this function $H_t(a)$ is concave in a at time $t = 0$.

Assumption 2a: An alternative concavity assumption on G: In developing some of our results, we find the following assumption that G is concave in an alternative direction useful. Assume that there is a point $(\{\bar{y}_{rt}(j)\}_{j\geq 1}, N_t^0(0))$ with state variables $\{\bar{N}_t(j)\}_{j\geq 1}$ in the domain of G such that

$$G(\{\bar{y}_{rt}(j)\}_{j\geq 1}, N_t^0(0); \{\bar{N}_t(j)\}_{j\geq 1}) = G_t^0, \tag{16}$$

where $G_t^0 = G(0, 0, \ldots, 0; \{\bar{N}_t(j)\}_{j\geq 1})$. Note that it may be the case that $N_t^0(0) < 0$; we only require that such a point be in the domain of G, not that it be feasible in the model. Let us use Y_{rt}^0 to denote the amount of the resource good used at this allocation

$$Y_{rt}^0 = \sum_{j\geq 1} \bar{y}_{rt}(j)\bar{N}_t(j) + \bar{y}_r(0)N_t^0(0) \tag{17}$$

Note as well that we may have $Y_{rt}^0 < 0$. This is allowed as long as such a point exists in the domain of G. It does not have to be feasible in the model to be defined here.

We then define $\widetilde{H}_t(a)$ over the domain $a \in [0, 1+\epsilon)$ for a fixed $\epsilon > 0$ as equal to the

value of G evaluated at the corresponding convex combination of the points $(\{\bar{y}_{rt}(j)\}_{j\geq 1}, \bar{N}_t(0))$ and $(\{\bar{y}_{rt}(j)\}_{j\geq 1}, N_t^0(0))$, with the state variables $\{\bar{N}_t(j)\}_{j\geq 1}$ equal to their baseline endogenous variables at time t. Assumption 2 is satisfied in examples 1-4. Assumption 2a is satisfied in example 1-3 while $N_t^0(0)$ is not defined in example 4. If example 4 were modified to have a linear entry margin as in examples 1-3, then Assumption 2a would be satisfied in this case as well.

Assumption 3: The state $\{N_t(j)\}_{j\geq 1}$ does not enter G: This assumption must be verified separately in each model. As we have seen in examples 1-3, this assumption is satisfied in simple versions of standard models that assume constant markups across firms (or, more generally, no correlation between markups and firms' relative productivities) and either no innovation expenditure by incumbents, or a certain form of identical and constant returns technologies for innovation by incumbent firms introduced by Klette and Kortum (2004).

Impact elasticity

We make use of the following summary measure of the response of the growth rate of aggregate productivity with respect to a policy induced change in production of the research good that we term the *impact elasticity of productivity growth with respect to output of the research good*. To construct such a measure, assume that as we shift smoothly from an original baseline path for innovation policies to the new path for innovation policies (parametrically taking a convex combination of the two policy paths) that the equilibrium path of the endogenous variables in the model moves in a differentiable manner from the baseline path of endogenous variables to the new path of endogenous variables. In this case, we approximate the contemporaneous change in the growth rate of aggregate productivity from period t to $t+1$ given a change in output of the research good at Y_{rt} as

$$\tilde{g}_{zt} - \bar{g}_z \approx \mathcal{E}_{gt}\left(\log \tilde{Y}_{rt} - \log \bar{Y}_r\right) \tag{18}$$

where

$$\mathcal{E}_{gt} = \sum_{j\geq 1} \frac{\partial G}{\partial y_r(j)} \bar{y}_{rt}(j) \frac{d\log y_{rt}(j)}{d\log Y_{rt}} + \frac{\partial G}{\partial N(0)} \bar{N}_t(0) \frac{d\log N_t(0)}{d\log Y_{rt}}. \tag{19}$$

Here the terms $d\log y_{rt}(j)$, $d\log N_t(0)$, and $d\log Y_{rt}$ are the rates of change of these equilibrium variables with respect to the parameter indexing the smooth shift from the baseline set of policies to the new policies evaluated at the baseline set of policies and, again, the partial derivatives of G are evaluated at the baseline path of the endogenous variables.

Note that in computing this impact elasticity, or impact effect, we are holding the current state variables $\{N_t(0)\}_{j\geq 1}$ fixed.

As is evident from equation (19), the exact value of this elasticity will depend, in general, on the specifics of the equilibrium responses of all of the endogenous variables to the specific policy change being modeled — that is, how is the change in the aggregate production of the research good is allocated across the different types of firms as we move from the baseline set of policies to the new policies. Without taking a particular stand on the specific policies being changed, we cannot characterize this endogenous change in the allocation of the research good across firms. Under assumption 1, however, we have the following lemma that this impact elasticity $\mathcal{E}_{gt}$ does not depend on the specifics of how the extra production of the research good is allocated across firms and hence does not depend on the specifics of the policies being changed.

Lemma 1. *If the baseline allocation is conditionally efficient (e.g satisfies Assumption 1), then the impact elasticity of productivity growth with respect to output of the research good, $\mathcal{E}_{gt}$, is independent of how the change in the output of the research good is allocated across incumbent and entering firms.*

Proof: This lemma is a simple application of the envelope theorem. From assumption 1, we have that the partial derivatives of G evaluated at the baseline allocation satisfy the first order necessary conditions of the Lagrangian formed from the problem of maximizing G subject to the resource constraint for the research good. These first order conditions are

$$\frac{\partial G}{\partial y_{rt}(j)}\bar{y}_{rt}(j) = \lambda_t \bar{y}_{rt}(j)\,\bar{N}_t(j) \tag{20}$$

and

$$\frac{\partial G}{\partial N(0)}\bar{N}_t(0) = \lambda_t \bar{y}_r(0)\,\bar{N}_t(0) \tag{21}$$

where λ_t is the Lagrange multiplier on the resource constraint for the research good (7) at time t.

Second, for any collection of variations in the endogenous variables and the total production of the research good that is feasible (i.e. satisfies constraint 7), we have

$$\sum_{j\geq 1}\bar{y}_{rt}(j)\bar{N}_t(j)\frac{d\log y_{rt}(j)}{d\log Y_{rt}} + \bar{y}_{rt}(0)\bar{N}_t(0)\frac{d\log N_t(0)}{d\log Y_{rt}} = \bar{Y}_{rt} \tag{22}$$

Plugging this result together with equations (20) and (21) into the definition of $\mathcal{E}_{gt}$, (19), gives

$$\mathcal{E}_{gt} = \lambda_t \bar{Y}_{rt} \tag{23}$$

This is a restatement of the usual result that the Lagrange multiplier measures the derivative of the objective with respect to a relaxation of the resource constraint regardless of how the extra resources are allocated across alternative uses.

We impose the assumption that along a BGP, the Lagrange multiplier λ_t (and hence the impact elasticity $\mathcal{E}_{gt}$) remains constant over time. This assumption can be verified directly in examples 1-3 above. In example 4, this assumption holds if N_t^m is constant over time on the initial BGP

Bounding the impact elasticity $\mathcal{E}_{gt}$ We now use assumption 1 together with assumptions 2 and/or 2a to derive an upper bound on the magnitude of the impact elasticities $\mathcal{E}_{gt}$ from the baseline growth rates of aggregate productivity $\bar{g}_z$ to which the model economy is calibrated. Note that assumption 3 is not required for this next result.

Proposition 1. *If assumptions 1 and 2 are satisfied, then the impact elasticity $\mathcal{E}_{gt}$ is bounded by*

$$\mathcal{E}_{gt} \leq \bar{g}_z - G_t^0 \tag{24}$$

where $\bar{g}_z$ is the growth rate of aggregate productivity on the baseline BGP and G_t^0 is the social depreciation rate of innovation expenditures. If assumptions 1 and 2a are satisfied, then the impact elasticity $\mathcal{E}_{gt}$ is bounded by

$$\mathcal{E}_{gt} \leq \left(\bar{g}_z - G_t^0\right)\left(\frac{\bar{Y}_r}{\bar{Y}_r - Y_{rt}^0}\right) \tag{25}$$

where Y_{rt}^0 is defined as in equation (17).

Proof. Consider the first bound in equation (24) derived from the definition of the function $H_t(a)$ in assumption 2. By this definition of $H_t(a)$, we have

$$H_t'(1) = \sum_{j\geq 1} \frac{\partial G}{\partial y_r(j)} \bar{y}_{rt}(j) + \frac{\partial G}{\partial N(0)} \bar{N}_t(0), \tag{26}$$

where the partial derivatives are evaluated at the baseline BGP allocation (with bars). By Lemma 1, we have

$$H_t'(1) = \lambda_t \left(\sum_{j\geq 1} \bar{y}_{rt}(j)\bar{N}_t(j) + \bar{y}_r(0)\bar{N}_t(0)\right) = \lambda_t \bar{Y}_r = \mathcal{E}_{gt} \tag{27}$$

By assumption 2 regarding the concavity of $H_t(a)$, we have $H_t'(1) \leq H_t(1) - H_t(0)$. Since $H_t(1) = \bar{g}_z$ and $H_t(0) = G_t^0$, this proves the first bound.

Now consider the second bound in equation (25) derived from the definition of the function $\widetilde{H}_t(a)$ in assumption 2a. By the definition of $\widetilde{H}_t(a)$, we have

$$\widetilde{H}_t'(1) = \frac{\partial G}{\partial N(0)}(\bar{N}_t(0) - N_t^0(0)), \tag{28}$$

where the partial derivative is evaluated at the baseline BGP allocation. By the first order condition (21) in Lemma 1, we have

$$\frac{\partial G}{\partial N\,(0)}\bar{N}_t\,(0)\,\frac{\bar{Y}_{rt}}{\bar{y}_r\,(0)\,\bar{N}_t\,(0)} = \lambda_t \bar{Y}_r = \mathcal{E}_{gt} \tag{29}$$

Putting these results together gives

$$\mathcal{E}_{gt} = \widetilde{H}_t'(1)\left(\frac{\bar{N}_t\,(0)}{\bar{N}_t\,(0) - N_t^0(0)}\right)\left(\frac{\bar{Y}_r}{\bar{y}_r\,(0)\,\bar{N}_t\,(0)}\right) \tag{30}$$

Using assumption 2a regarding the concavity of $\widetilde{H}_t(a)$ gives $\widetilde{H}_t'(1) \leq \widetilde{H}_t(1) - \widetilde{H}_t(0)$. By construction $\widetilde{H}_t(1) = \bar{g}_z$ and $\widetilde{H}_t(0) = G_t^0$. Thus, using the fact that $\left(\bar{N}_t\,(0) - N_t^0(0)\right)\bar{y}_r\,(0) = \bar{Y}_r - Y_{rt}^0$, we have

$$\mathcal{E}_{gt} \leq \left(\bar{g}_z - G_t^0\right)\left(\frac{\bar{Y}_r}{\bar{Y}_r - Y_{rt}^0}\right) \tag{31}$$

This proves the result. □

Social depreciation of innovation expenditures and the Bound on the Impact Elasticity of Productivity Growth with respect to aggregate production of the research good A key implication of Proposition 1 is that we are able to derive an upper bound on the impact elasticity of the growth of aggregate productivity with respect to the output of the research good, $\mathcal{E}_{gt}$, that depends on a small number of sufficient statistics. Consider first the bound given by (24). If there is no social depreciation of innovation expenditures (i.e. $G_t^0 = 0$), then the impact elasticity is simply bounded by the initial calibrated growth rate of aggregate productivity $\mathcal{E}_{gt} \leq \bar{g}_z$. This bound is quite restrictive if the baseline growth rate of productivity to which the model is calibrated is low. In contrast, if there is social depreciation (i.e. $G_t^0 < 0$) then the bound on the impact elasticity is looser.

Innovation technologies for incumbent and entering firms Consider next the bound (25). This bound is equal the first bound implied by (24) multiplied by the term $\bar{Y}_{rt}/\left(\bar{Y}_{rt} - Y_{rt}^0\right)$. If there is no innovation by incumbent firms as in examples 1 and 2, this term is equal to 1, and hence the two bounds are equal, as we discuss below. When there is innovation

by incumbent firms, then the magnitude of this term is determined by the differences in the average cost of innovation for incumbent and entering firms as follows. To see this, consider the ratio

$$\frac{\bar{Y}_r - 0}{\bar{g}_z - G_t^0}. \tag{32}$$

This is the cost savings in terms of the research good of reducing the growth rate of aggregate productivity from $\bar{g}_z$ to G_t^0 by decreasing all incumbent and entering firms' research expenditures proportionally from the baseline values to zero (this is the variation considered in the definition of $H_t(a)$ in assumption 2). Consider now the ratio

$$\frac{\bar{Y}_{rt} - Y_{rt}^0}{\bar{g}_z - G_t^0}. \tag{33}$$

This is the cost savings in terms of the research good of reducing the growth rate of aggregate productivity from $\bar{g}_z$ to G_t^0 by decreasing *only* entering firms' research expenditures from the baseline value to the required value such that the growth rate is G_t^0 (this is the variation considered in the definition of $H_t(a)$ in assumption 2a). If the average cost implied by (32) is lower than the average cost implied by (33), then the term $\bar{Y}_r / \left(\bar{Y}_r - Y_{rt}^0\right)$ is less than one and the bound implied by (25) is tighter than the bound implied by (24). In example 3, we show that this is the case when incumbents have a lower average cost of innovation than entrants at the baseline BGP allocations. More generally, the term $\bar{Y}_{rt} / \left(\bar{Y}_{rt} - Y_{rt}^0\right)$ corresponds to the gap between the average cost (in social terms) of innovation using incumbents' technologies and using entrants' technologies for innovation.

Application of bounds to model examples We now illustrate the application of the bounds on the impact elasticity of aggregate productivity growth with respect to changes in aggregate production of the research good. A comparison of examples 1, 2, and 3 highlights the role of assumptions regarding the social depreciation of innovation expenditures and the shapes of the technologies for innovation by incumbent and entering firms.

Impact Elasticity in Example 1: In the simple Quality Ladders model of example 1, assumptions 1 - 2 can be verified immediately. The exact impact elasticity of aggregate productivity growth with respect to aggregate production of the research good can be calculated directly as

$$\mathcal{E}_{gt}^{QL} = \frac{1}{\rho - 1} \frac{\exp((\rho - 1)\bar{g}_z) - 1}{\exp((\rho - 1)\bar{g}_z)}$$

and it is straightforward to verify directly that exact elasticity converges to $\bar{g}_z$ as $\rho \to 1$

and to zero as $\rho \to \infty$. Thus, since in this model $G_t^0 = 0$, the bound in (24) becomes tight as $\rho \to 1$. Assuming a higher elasticity of substitution between intermediate goods simply reduces the exact impact elasticity below the bound.

Note as well that this model also satisfies assumption 2a and thus it is also possible to apply the bound in (25). Since innovation is done only by entering firms, however, we have that $Y_{rt}^0 = 0$ and that this bound is also equal to $\bar{g}_z$.

This example highlights the role that the assumption that there is no social depreciation of innovation expenditures (i.e. that $G_t^0 = 0$) plays in restricting the quantitative implications of the model for the impact elasticity $\mathcal{E}_{gt}^{QL}$.

Impact Elasticity in Example 2: Given the definition of G in the Expanding Varieties model example 2, assumptions 1-2 can again be verified immediately. The exact impact elasticity of aggregate productivity growth with respect to aggregate production of the research good in this example can be calculated directly as

$$\mathcal{E}_{gt}^{EV} = \frac{1}{\rho - 1} \frac{\exp((\rho - 1)\bar{g}_z) - \exp((\rho - 1)G^0)}{\exp((\rho - 1)\bar{g}_z)},$$

where in this case,

$$G_t^0 = G^0 = \frac{1}{\rho - 1} \log((1 - \delta_f) \exp((\rho - 1)\Delta_z)).$$

Clearly, we see by direct comparison that the exact elasticity in the Expanding Varieties model exceeds that in the Quality Ladders model given the same calibrated values of ρ and $\bar{g}_z$ as long as $(1 - \delta_f) \exp((\rho - 1)\Delta_z) < 1$. This relationship between the exact impact elasticities computed in these two example models reflects the role of social depreciation of innovation expenditures in the two models. That is, if $(1 - \delta_f) \exp((\rho - 1)\Delta_z) < 1$, we also have $G^0 < 0$, and hence the bound in (24) is also looser in the Expanding Varieties model compared to the Quality Ladders model. Again, this bound on the exact elasticity becomes tight as $\rho \to 1$.

This model satisfies assumption 2a and thus it is also possible to apply the bound in (25). Since innovation is done only by entering firms, however, we have that $Y_{rt}^0 = 0$ and hence both approaches yield the same bound on the impact elasticity of aggregate productivity in the Expanding Varieties model.

Impact Elasticity in Example 3: We now turn to the simple version of the Klette Kortum model with innovation by both incumbents and entrants. To satisfy assumption 1 regarding the conditional efficiency of the baseline allocation, from equations (20) and

(21), we have that the baseline allocation must satisfy $d'(\bar{y}_r(1)) = 1/\bar{y}_r(0)$.[19] With this assumption, we have that the exact impact elasticity for the model in example 3 is given by

$$\mathcal{E}_{gt}^{KK} = \frac{1}{\rho-1}\frac{\exp((\rho-1)\bar{g}_z)-1}{\exp((\rho-1)\bar{g}_z)}\frac{d'(\bar{y}_r(1))\bar{y}_r(1)+\bar{N}(0)}{d(\bar{y}_r(1))+\bar{N}(0)} = \frac{1}{\rho-1}\frac{\exp((\rho-1)\bar{g}_z)-1}{\exp((\rho-1)\bar{g}_z)}\frac{\bar{Y}_r}{\bar{Y}_r-Y_r^0}$$

where the last equality follows from conditional efficiency and the definition of $Y_{rt}^0 = y_{rt}(1) - d(y_{rt}(1))\bar{y}_r(0)$ in equation (17).

This example shares with the Quality Ladders model the assumption that there is no social depreciation of innovation expenditures in the sense that $G_t^0 = 0$, and hence the bound on the impact elasticity given in expression (24) is the same as in the Quality Ladders model $\mathcal{E}_{gt}^{KK} \leq \bar{g}_z$. But notice that the true elasticity in this model is lower than in the Quality Ladders model whenever $\bar{Y}_r/(\bar{Y}_r - Y_r^0) < 1$ or, equivalently, $d'(\bar{y}_r(1))\bar{y}_r(1) < d(\bar{y}_r(1))$. This inequality holds when the average cost of producing an innovation with the incumbent firms' innovation technology is lower than the average cost of producing an innovation with the entrants' technology in the baseline conditionally efficient allocation. If $d(0) = 0$ and $d(.)$ is strictly concave, then this condition will be satisfied in any conditionally efficient allocation in which incumbents innovate. We see this difference between the two models reflected in the second bound given in equation (31) as this bound includes the same term $\bar{Y}_r/(\bar{Y}_r - Y_r^0)$ that appears in the true elasticity. This second bound becomes tight as $\rho \to 1$.

Impact Elasticity in Example 4: Our proposition 1 applies to the version of the Klette-Kortum model studied in Lentz and Mortensen (2014) under the assumptions stated in that proposition including the assumption that the initial allocation is conditionally efficient. In contrast, if the initial allocation is not conditionally efficient, then this proposition does not apply. We now turn to a discussion of the impact elasticity of aggregate productivity growth with respect to a change in aggregate production of the research good when assumption 1 is not satisfied.

Lack of conditional efficiency and the impact elasticity In models with multiple types of firms engaged in innovative activity (e.g. examples 3 and 4), it is possible, and even likely, that the baseline allocation on an initial BGP is not conditionally efficient. In this case, our Lemma 1 fails to hold, and the magnitude of the more general impact elasticity given in equation (19) depends on the specific details of how the additional units of

[19] In the appendix we show that this condition is satisfied in equilibrium in the baseline BGP if entrants and incumbents are subsidized at the same rate.

the research good are allocated across different types of firms in response to a specific policy change. Note, however, that if Assumption 2 is satisfied, then the bound on the impact elasticity developed in Proposition 1 applies to changes in aggregate output of the research good that are allocated proportionally across all incumbent and entering firms (e.g. $\frac{d \log y_{rt}(j)}{d \log Y_{rt}} = \frac{d \log N_t(0)}{d \log Y_{rt}}$ for all j). Likewise, if Assumption 2a is satisfied, then the corresponding bound developed in that Proposition applies to changes in aggregate output of the research good that are allocated entirely to entering firms. What happens when Assumption 1 is violated is that reallocations of innovation spending that raise spending for some firms and lowers it for others may have first-order effects on the growth of aggregate productivity.

For example, in the simple Klette-Kortum model of example 3, if we assume that $d'(\bar{y}_r(1)) \neq 1/\bar{y}_r(0)$ in the initial BGP, then, the impact elasticity defined in (19) can be greater or smaller than the impact elasticity we calculated above depending on whether the reallocation of the research good induced by the policy change favors or disfavors the firms with the higher marginal contribution to aggregate productivity growth. Moreover, the approach we took to bounding this elasticity in Proposition 1 cannot be applied because this impact elasticity may not be well defined — if conditional efficiency fails, it is theoretically possible to reallocate the research good across firms to increase the growth rate of aggregate productivity without increasing the aggregate production of the research good at all.

While we cannot make analytical statements about the impact elasticity of aggregate productivity growth with respect to changes in aggregate production of the research good in this case, we can, however examine the potential additional welfare gains that might arise from moving from a conditionally inefficient allocation to a conditionally efficient allocation along a BGP with a fixed growth rate. If the initial allocation of the research good across firms on the initial baseline BGP at time t is not conditionally efficient, then it is possible to attain the same BGP growth rate, $\bar{g}_z$, with a lower aggregate output of the research good, Y^*_{rt}, which solves the problem:

$$Y^*_{rt} = \min_{\{y_r(j)\}_{j\geq 1}, N(0)} \sum_{j\geq 1} y_r(j) N_t(j) + \bar{y}_r(0) N(0)$$

subject to

$$\bar{g}_z = G\left(\{y_r(j)\}_{j\geq 1}, N(0); \{N_t(j)\}_{j\geq 1}\right).$$

If the initial allocation of the research good is not conditionally efficient, then $Y^*_{rt} < \bar{Y}_r$ and it is possible to maintain the same growth rate, $\bar{g}_z$, and immediately increase consump-

tion to a higher level by increasing labor in the production of intermediate goods from $\bar{L}_p = \frac{Z_t^{\gamma-1}\bar{Y}_r}{A_{rt}}$ to $L_{pt}^* = 1 - \frac{Z_t^{\gamma-1}Y_{rt}^*}{A_{rt}}$. Under assumption 3, the solution to this problem does not depend on the state vector $\{N_t(j)\}_{j\geq 1}$ describing the distribution of incumbent firms across types and hence is independent of time. In this case, if the initial allocation of the research good is not conditionally efficient, then $Y_{rt}^* < \bar{Y}_r$ and it is possible to maintain the same growth rate, $\bar{g}_z$, and immediately increase consumption to a new permanently higher level by increasing labor in the production of intermediate goods from $\bar{L}_p = \frac{Z_t^{\gamma-1}\bar{Y}_r}{A_{rt}}$ to $L_p^* = 1 - \frac{Z_t^{\gamma-1}Y_r^*}{A_{rt}}$.

This reasoning suggests that one might approximate the welfare gains to be obtained from innovation policies starting from an initial allocation that is not conditionally efficient in two steps. First, one can compute the permanent increase in consumption that can be achieved by moving to a conditionally efficient allocation with the same growth rate of aggregate productivity as on the initial BGP, as described above. Then second, one can compute using our baseline analytic results the aggregate implications of using innovation policies to alter the time path for the innovation intensity of the economy and the growth rate of aggregate productivity and output starting from a new baseline allocation that is conditionally efficient. We illustrate such a calculation for our model in example 3 in the next section. This exercise requires one to fully parameterize the function G beyond the sufficient statistics discussed above.[20]

Dynamics of aggregate productivity and GDP

So far, we have characterized the elasticity on impact of aggregate productivity growth with respect to changes in aggregate production of the research good. We now consider how one can use these results together with assumption 3 to characterize the full transition dynamics of aggregate productivity and GDP to a change in the innovation intensity of the economy. The following Proposition characterizes the dynamics of aggregate productivity to a change in innovation policies.

[20]The paper by Garcia-Macia et al. (2015) is an interesting study of a version of the Klette and Kortum (2004) model modified to include innovations by incumbents on their own products and the possibility for an increasing number of products in the aggregate. They show that one can match a rich set of data on firm and plant dynamics in this model by setting innovation rates for the different types of innovations that can occur directly as parameters of the model to be estimated. As we discuss in the Appendix, the features of equilibrium that these authors focus on such as the extent of business stealing are important inputs in designing innovation policies to implement the optimal level and allocation of innovative spending across firms. However, since these authors do not specify the technology G linking investments in innovation by firms to growth, their analysis does not permit identification of the impact effect and transition dynamics following a change in the innovation intensity of the economy or whether the baseline equilibrium allocation in the data deviates from conditional efficiency and, if so, in which direction.

Proposition 2. *Consider an economy that satisfies Assumptions 1 and 3. Suppose that economy is on an initial baseline BGP and, at time $t = 0$, a change in innovation policies induces a new path for the innovation intensity of the economy given by $\{s'_{rt}\}_{t=0}^{\infty}$. Assume that average markups μ_t remain constant on the new transition path. Then the new path for aggregate productivity $\{Z'_t\}_{t=1}^{\infty}$ to a first-order approximation is given by*

$$\log Z'_{t+1} - \log \bar{Z}_{t+1} = \sum_{k=0}^{t-1} \Gamma_k \left(\log s'_{rt-k} - \log \bar{s}_r\right) \tag{34}$$

where $\bar{Z}_t = \exp(t\bar{g}_z)\bar{Z}_0$, with $\Gamma_0 = \bar{L}_p \mathcal{E}_{g0}$ denoting the impact elasticity of aggregate productivity growth from period t=0 to t=1 in response to the change in the innovation intensity of the economy at time $t = 0$, and

$$\Gamma_{k+1} = \left[1 - (1-\gamma)\mathcal{E}_{g0}\right] \Gamma_k \textit{ for } k \geq 0. \tag{35}$$

Proof. We prove this result by calculating two key elasticities in our model. The first key elasticity is the elasticity of research output Y_r with respect to a change in the innovation intensity of the economy s_r and a change in the path of aggregate productivity Z along the transition. The second key elasticity is the impact elasticity of aggregate productivity growth with respect to changes in the production of the research good Y_{rt} characterized in Lemma 1.

To calculate the elasticity of research output Y_r with respect to a change in the innovation intensity of the economy s_r and a change in the path of aggregate productivity Z along the transition, note that from equations (8) and (15) and assuming that average markup μ_t is constant over time, we have that, to a first-order approximation,

$$\log Y'_{rt} - \log \bar{Y}_r = \bar{L}_p \left(\log s'_{rt} - \log \bar{s}_r\right) - (1-\gamma)\left(\log Z'_t - \log \bar{Z}_t\right). \tag{36}$$

That the coefficient $\bar{L}_p$ is constant over time follows from the assumption that the initial allocation is on a BGP. That there is no term her reflecting the impact of changes in the average markup on the allocation of labor between production and research follows from the assumption that average markups are constant on the new transition path.

Now turn to the second elasticity. By Lemma 1 and assumption 3, we have

$$\log Z'_{t+1} - \log Z'_t \approx \bar{g}_z + \mathcal{E}_{gt}\left(\log Y'_{rt} - \log \bar{Y}_r\right). \tag{37}$$

In a baseline BGP, $\mathcal{E}_{gt}$ is assumed to be constant at $\mathcal{E}_{g0}$, so this equation is equivalent to

$$\log Z'_{t+1} - \log \bar{Z}_{t+1} \approx \log Z'_t - \log \bar{Z}_t + \mathcal{E}_{g0} \left(\log Y'_{rt} - \log \bar{Y}_r \right). \tag{38}$$

Plugging in (36) and (37) into (38) gives

$$\log Z'_{t+1} - \log \bar{Z}_{t+1} = \left[1 - (1-\gamma)\mathcal{E}_{g0} \right] \left(\log Z'_t - \log \bar{Z}_t \right) \tag{39}$$

$$+\Gamma_0 \left(\log s'_{rt} - \log \bar{s}_r \right),$$

which can be used to derive expression (34). □

Proposition 2 gives us an analytical expression for the dynamics of aggregate productivity in the transition to a new BGP following an unanticipated and permanent change in innovation policies as a function of the transition path for the innovation intensity of the economy that arises in equilibrium as a result of that change in policies. From this proposition, we can compute the long-run change in aggregate productivity that corresponds to a given permanent change in innovation intensity of the economy as follows.

Corollary 1. *Consider a permanent change in innovation policies. Assume that the economy converges to a new BGP with innovation intensity s'_r. Under the assumptions in Proposition 2, the gap in aggregate productivity between the old and new BGP converges, to a first-order approximation, to*

$$\log Z'_t - \log \bar{Z}_t = \frac{\bar{L}_p}{1-\gamma} \left(\log s'_r - \log \bar{s}_r \right) \tag{40}$$

in the semi-endogenous growth case ($\gamma < 1$). In the endogenous growth case ($\gamma = 1$) the gap in aggregate productivity between the old and new BGP is unbounded. The new growth rate of aggregate productivity to a first-order approximation is given by

$$\log Z'_{t+1} - \log Z'_t = \bar{g}_z + \Gamma_0 \left(\log s'_{rt} - \log \bar{s}_r \right). \tag{41}$$

Proof. To derive expression (40), we use the fact that $\frac{\bar{L}_p}{1-\gamma}$ is equal to $\sum_{k=0}^{\infty} \Gamma_k$ in expression (34) if that sum converges. With endogenous growth, $\gamma = 1$, and hence $\Gamma_k = \Gamma_0$ for all k. Equation (41) follows from taking the first difference of equation (34). □

We next derive the transition path for GDP as a function of the transition paths for aggregate productivity, the innovation intensity of the economy, and the equilibrium rental rate on physical capital.[21]

[21] The magnitude of the change in the rental rate of physical capital in the transition is related to the

Corollary 2. *Under the assumptions in Proposition 2, the path of GDP corresponding to the policy experiment in Proposition 2 is given, to a first-order approximation, by*

$$\log GDP'_t - \log G\bar{D}P_t = \frac{1}{1-\alpha}\left(\log Z'_t - \log \bar{Z}_t\right) - \bar{L}_r\left(\log s'_{rt} - \log \bar{s}_r\right) - \frac{\alpha}{1-\alpha}\left(\log R'_{kt} - \log \bar{R}_k\right) \tag{42}$$

under the old measurement system in which innovation is expensed, and is given by the above plus $\frac{\bar{s}_r}{1+\bar{s}_r}\left(\log s'_{rt} - \log \bar{s}_r\right)$ *under a measurement system in which all expenditures on innovation are included in measured GDP.*

Proof. We prove this result by taking the log of GDP, which under the old measurement system is equal to the log of Y:

$$\log Y'_t - \log \bar{Y}_t = \log Z'_t - \log \bar{Z}_t + \alpha(\log K'_t - \log \bar{K}_t) + (1-\alpha)\left(\log L'_{pt} - \log \bar{L}_p\right)$$

Using $R_{kt} = \frac{\alpha}{\mu}\left(1+\tau_s\right)\frac{Y_t}{K_t}$ and equation (15) we obtain the expression above. To derive the result under the new measure of GDP, we must add in expenditures on research $P_{rt}Y_{rt}$, which we can do by multiplying the old measure of GDP by $(1+s_{rt})$. □

Given the result in Corollary 2, it is straightforward to calculate the response of GDP in the long run. Since, with semi-endogenous growth, in the long run the interest rate and rental rate on physical capital return to their levels on the initial BGP (i.e $\lim_{t\to\infty} \log R'_{kt} - \log \bar{R}_k = 0$), the long run response of GDP is a simple function of the long run responses of aggregate productivity and the innovation intensity of the economy. In particular, from Corollary 1, we have that with $\gamma < 1$, the long run response of GDP is given by

$$\lim_{t\to\infty} \log GDP'_t - \log G\bar{D}P_t = \left[\frac{1}{1-\alpha}\frac{\bar{L}_p}{1-\gamma} - \bar{L}_r\right]\lim_{t\to\infty}\left(\log s'_{rt} - \log \bar{s}_r\right).$$

The relationship between the impact elasticity $\mathcal{E}_{g0}$ and the dynamics of productivity From Corollary 1, we have that in the semi-endogenous growth case, the long run elasticity of aggregate productivity with respect to a permanent change in the innovation intensity of the economy given in equation (40) is independent of the impact elasticity $\mathcal{E}_{g0}$. This impact elasticity does, however, affect the model's transition dynamics from the initial BGP to the new BGP. We illustrate this role of the impact elasticity in shaping the transition dynamics by comparing the transition dynamics in the models of examples 1 and 3.

equilibrium transition path for the interest rate. From the Euler equation for physical capital, $\log R'_{kt} - \log \bar{R}_k = \frac{\bar{r}}{d_k+\bar{r}}\log\left(\bar{r}_{t-1}/\bar{r}\right)$ for $t \geq 1$. Solving for the path of the interest rate requires fully solving for the model transition, which we do in the Appendix.

The models in examples 1 and 3 are both based on the Quality Ladders framework, but the Klette-Kortum style model in example 3 adds innovation by incumbent firms using a concave innovation technology $d(y_r)$. We have seen that this assumption reduces the impact elasticity of aggregate productivity growth with respect to a policy induced change in production of the research good. We now examine the impact of this assumption on the implications of this model for the transition dynamics of aggregate productivity following a permanent policy induced change in the innovation intensity of the economy. We do so by comparing the implications of the two models for the coefficients $\{\Gamma_k\}$ that determine the elasticities of aggregate productivity and GDP with respect to changes in the innovation intensity of the economy. We do so in the following Proposition.

Proposition 3. *Consider the Klette-Kortum style model in example 3 and the Quality Ladders model in example 1. Assume that they are calibrated to share the same parameters γ and ρ and to have the same implications for $\bar{g}_z$ and $\bar{L}_p$ on the initial BGP. Then the elasticities $\left\{\Gamma_k^{KK}\right\}$ in the Klette-Kortum version of the model are related to the elasticities $\left\{\Gamma_k^{QL}\right\}$ in the Quality Ladders version of the model defined in Proposition 2 as follows. For any $T \geq 1$,*

$$\sum_{k=0}^{T} \Gamma_k^{KK} \leq \sum_{k=0}^{T} \Gamma_k^{QL}.$$

With semi-endogenous growth, $\gamma < 1$, as $T \to \infty$,

$$\sum_{k=0}^{\infty} \Gamma_k^{KK} = \sum_{k=0}^{\infty} \Gamma_k^{QL}$$

while with endogenous growth ($\gamma = 1$), $\Gamma_k^{KK} = \Gamma_0^{KK} \leq \Gamma_0^{QL} = \Gamma_k^{QL}$ for all $k \geq 0$.

Proof. For both models, we have,

$$\sum_{k=0}^{T-1} \Gamma_k = \frac{\bar{L}_p}{1-\gamma}\left[1 - \left(1 - (1-\gamma)\mathcal{E}_{g0}\right)^T\right].$$

The result follows from the observation that $\mathcal{E}_{g0}^{KK} \leq \mathcal{E}_{g0}^{QL}$. □

What this result implies is that if we consider innovation policy changes in the Klette-Kortum model that produce the same transition path for the innovation intensity of the economy $\{s'_{rt}\}$ as the innovation policy changes considered in the Quality Ladders version of the model, then the Quality Ladders version of the model will imply a larger change in aggregate productivity up to any period T along the transition path, to a first-order approximation. In the long-run, as $T \to \infty$, with semi-endogenous growth, the two

models deliver the same response of aggregate productivity and GDP. Likewise, with endogenous growth, for every T, the response of the growth rate will be larger in the Quality Ladders version of the model than in the Klette-Kortum version of the model, again to a first-order approximation. A similar result can be derived comparing the transition dynamics implied by Expanding Varieties model without and with innovation by incumbents (as in Atkeson and Burstein 2010).

Aggregate dynamics and optimal innovation intensity Expression (42) in Corollary 2 is useful for understanding the transition dynamics of measured GDP (corresponding to the resources available for consumption and physical investment) in models that satisfy assumptions 1 and 3. With a permanent increase in the innovation intensity of the economy, GDP falls initially as labor is reallocated from current production to research, and then grows in the transition as the impact of a permanent change in the innovation intensity of the economy on aggregate productivity cumulates over time.

The impact on welfare of a permanent increase in the innovation intensity of the economy clearly depends on the trade-off between the short and long run changes in GDP that result from this increase in innovation expenditures as captured by the parameters Γ_k. To gain intuition for this welfare tradeoff, consider a variation in innovation policies that raises the innovation intensity of the economy only in period $t = 0$, so $s'_{r0} > \bar{s}_r$ and in all other time periods $t \neq 0$, $s'_{rt} = \bar{s}_r$. From Corollary 2, if we ignore changes in the rental rate on physical capital, we have that the log of resources available for consumption and physical investment (i.e. GDP) falls in period $t = 0$ by $\bar{L}_r \left(\log s'_0 - \log \bar{s}_r\right)$ and then rises in every period $t \geq 1$ by $\frac{\Gamma_{t-1}}{1-\alpha} \left(\log s'_0 - \log \bar{s}_r\right)$. In the appendix, we show that on the optimal BGP allocation, this perturbation of the path of innovation expenditures has no first order impact on welfare, which is equivalent to the condition that the socially optimal BGP allocation must satisfy

$$\left[\sum_{k=0}^{\infty} \tilde{\beta}^{1+k} \frac{\Gamma_k}{1-\alpha} - \bar{L}_r\right] \left(\log s'_0 - \log \bar{s}_r\right) = 0$$

for small perturbations of $\bar{s}_r$, where $\tilde{\beta} = \beta \exp\left((1-\xi)\,\bar{g}_y\right)$ is equal to the ratio between the growth rate and the interest rate in the BGP. This condition implies that on the optimal BGP allocation,[22]

$$s_r^* = (1-\alpha)\frac{L_r^*}{L_p^*} = \frac{\tilde{\beta}\mathcal{E}_g^*}{1-\tilde{\beta}\left[1-(1-\gamma)\mathcal{E}_g^*\right]}, \tag{43}$$

[22]One can also obtain this expression using the variational argument proposed by Jones and Williams (1998).

where s_r^*, L_r^*, and L_p^* are the optimal BGP levels of these variables and $\mathcal{E}_g^*$ is the impact elasticity of aggregate productivity growth with respect to changes in aggregate production of the research good evaluated on the optimal BGP. Hence, the optimal innovation intensity of the economy is solely determined by these variables, together with $\tilde{\beta}$ and γ. Note that if $\gamma < 1$ so that the model has semi-endogenous growth, then the impact elasticity $\mathcal{E}_g^*$ is bounded above by the exogenous BGP productivity growth rate, so we can bound s_r^* from above given parameters $\tilde{\beta}$, γ and the exogenous BGP productivity growth. In the appendix we derive the uniform innovation subsidy and production subsidy to implement the optimal allocation as a function of the parameters of the model (including equilibrium variables such as the markup).

Models that violate Assumption 3 and transition dynamics How would our results be altered for models that do not satisfy our assumption 3 that the state vector $\{N_t(j)\}_{j\geq 1}$ describing the distribution of incumbent firms across types does not enter the function G? Our Proposition 1 does not rely on our assumption 3. Hence our results regarding the impact elasticity of aggregate productivity growth with respect to a change in aggregate production of the research good holds in models in which this state variable does enter into the function G such as the one considered in examples 3 and 4 if different incumbent firms are offered different innovation subsidies and hence do they not all choose $y_{rt}(j) = n(j)y_{rt}(1)$. In contrast, our assumption 3 is required for the Proposition 2 describing the dynamics of aggregate productivity in the transition following a change in the equilibrium path for aggregate output of the research good simply because, in this case, the growth rate of aggregate productivity is altered, to a first order approximation, along the new equilibrium path of this state vector by

$$g'_{zt} - \bar{g}_z \approx \sum_{j\geq 1} \frac{\partial G}{\partial y_r(j)} \bar{y}_{rt}(j) \left(\log y'_{rt}(j) - \log \bar{y}_{rt}(j)\right) + \frac{\partial G}{\partial N(0)} \bar{N}_t(0) \left(\log N'_t(0) - \log \bar{N}_t(0)\right) +$$

$$\sum_{j\geq 1} \frac{\partial G}{\partial N(j)} \left(N'_t(j) - \bar{N}_t(j)\right).$$

Under our assumption 3, the last set of terms in this approximation is zero and under our assumption 1 (conditional efficiency), the remaining terms sum to the initial impact elasticity $\mathcal{E}_g$ at every date. Without assumption 3, we can no longer study this last set of terms analytically and, thus we no longer have the simple relationship between the coefficients Γ_0 and Γ_k that we derived in Proposition 2. What remains to be seen in quantitative work is whether consideration of these extra terms in estimated models lead to substan-

tially different quantitative results than what would be found in simpler models along the lines of those in examples 1 and 2 calibrated similarly in terms of their parameters for ρ and γ and their implications for $\bar{g}_z$, $\bar{L}_p$, and the social rate of depreciation. We turn to this question in the next section.

5 Quantitative implications of analytical results

In this section, we illustrate with particular numerical examples the use of our simple analytical results for quantitative analysis of the aggregate implications of changes in innovation policies in our model. In particular, we focus on illustrating the model parameters that are most important in shaping the implications of our model for the response of aggregate productivity over the medium term (20 years) and long term to innovation policies that have a permanent impact of the innovation intensity of the economy. We use our four example economies to illustrate these key model features. Additional details of our baseline calibration (based on model example 3) are presented in the Appendix.

Model Implied Dynamics in Example 1: Consider first a calibration of our simple Quality Ladders model from example 1 in which the time period is set to one year, the elasticity of substitution between goods is $\rho = 4$, and the growth rate of aggregate productivity on the initial BGP is $\bar{g}_z = .0125$, or 1.25%, so that (for our choice of $\alpha = 0.37$) the growth rate of output per worker is 2%, similar to that experienced in the U.S. over the post-war period.[23] With these assumptions, the exact impact elasticity of aggregate productivity growth with respect to a change in aggregate production of the research good is $\mathcal{E}^{QL}_{g0} = .0122$ which is very close to its maximum value of $\bar{g}_z = .0125$. Assume that the initial innovation intensity of the economy is $\bar{s}_r = .11$, similar to the levels estimated by Corrado et al. (2009) for the United States over the last few years. Using equation (15) and calibrated values for the average markup $\mu = 1.136$ and the physical capital share α, we have a share of labor employed in current production of $\bar{L}_p = .83$. With these assumptions, the initial elasticity of aggregate productivity with respect to changes in the innovation intensity of the economy is given by $\Gamma_0 = .01$. This value of Γ_0 is close to its largest feasible value equal to $\bar{g}_z = .0125$. Hence, we regard this numerical example as useful in illustrating close to a best case scenario for finding large elasticities of aggregate productivity to changes in the innovation intensity of the economy and high sensitivity

[23] Note that if one does a full growth accounting exercise including human capital and investment specific technical change, one gets an even smaller value for $\bar{g}_z$. For example, Jorgenson et al. (2014), estimates growth of TFP between 1947 and 2010 of roughly $\bar{g}_z = .007$.

of these responses to the degree of intertemporal knowledge spillovers.

The only other parameter of the model that we need to specify to apply our analytical results is the intertemporal knowledge spillover parameter γ which, together with $\mathcal{E}_{g0}^{QL}$, determines the decay rate of the elasticities Γ_k. We consider three alternative values of the intertemporal knowledge spillover parameter γ. The first case we call the *high spillover* case. In it we set $\gamma \to 1$. This corresponds to the limiting endogenous growth case. The second case we call the *medium spillover* case, and in it we set $\gamma = 0$. The third case we call the *low spillover* case, and in it we set $\gamma = -2$.

Consider an unanticipated and permanent increase in innovation subsidies that immediately and permanently raise the innovation intensity of the economy to $s_r' = \bar{s}_r + .03 = .14$, so that $\log s_r' - \log \bar{s}_r = .24$. In the appendix, we show that in our model examples, changes in innovation subsidies uniformly applied to entering and incumbent firms result in the long-run in a change in fiscal expenditures relative to GDP of $\frac{E'}{GDP'} - \frac{\bar{E}}{\bar{GDP}} = s_r' - \bar{s}_r$. Therefore, the policy-induced increase in the innovation intensity of the economy would require a recurring fiscal expenditure on innovation subsidies equal to 3 percent of GDP in the long-run, roughly equal to the total revenue collected from corporate profit taxes relative to GDP in 2007 in the U.S. In this sense, we regard this as a large change in the innovation intensity of the economy.

We illustrate our first order approximation to the dynamics of the transition for our model economy in Figure 1. In constructing this figure, as shown in Panel A, we assume that on this transition path, the physical capital to output ratio is constant at its BGP level. As shown in Panel B, we also assume that that the innovation intensity of the economy jumps to its new BGP level immediately, i.e. $\log s_{rt}' - \log \bar{s}_r = .24$ for all $t \geq 0$. We do so to illustrate the quantitative implications of the model given a specified path for the innovation intensity of the economy. In the appendix, we compute the full model (using example 3) to examine the actual transition path for the innovation intensity and physical capital to output ratio of the economy given our specific policy experiment to evaluate the usefulness of these approximations. We now use this example to discuss four main quantitative implications of our analytical results.

(i) Intertemporal knowledge spillovers and long run elasticities In the bottom left and right panels of Figure 1 (panels E and F) we show the transition paths for aggregate productivity and GDP, both as ratios to their levels on the initial BGP for the first 100 years of the transition. From Corollary 1, we have that the response of aggregate productivity in the long-run is very sensitive to the choice of intertemporal knowledge spillover parameter γ. As $\gamma \to 1$, the response of aggregate productivity becomes infinite because

the growth rate of aggregate productivity is approximately $\Gamma_0(\log s_r' - \log \bar{s}_r) = 0.0024$ (24 basis points) above its initial level permanently (in the limit). In contrast, if $\gamma = 0$, then the long run response of aggregate productivity relative to its initial BGP path is $\bar{L}_p \times .24 = .2$ so productivity is up only 20% relative to its level on the initial BGP in the long run. If $\gamma = -2$, then aggregate productivity rises only 6.7% relative to its initial BGP path in the long run. Using the formula for the dynamics of GDP under the assumption that the physical capital to output ratio remains constant obtained in Corollary 2, we see that the response of GDP in the long run is also very sensitive to the choice of intertemporal knowledge spillover parameter.

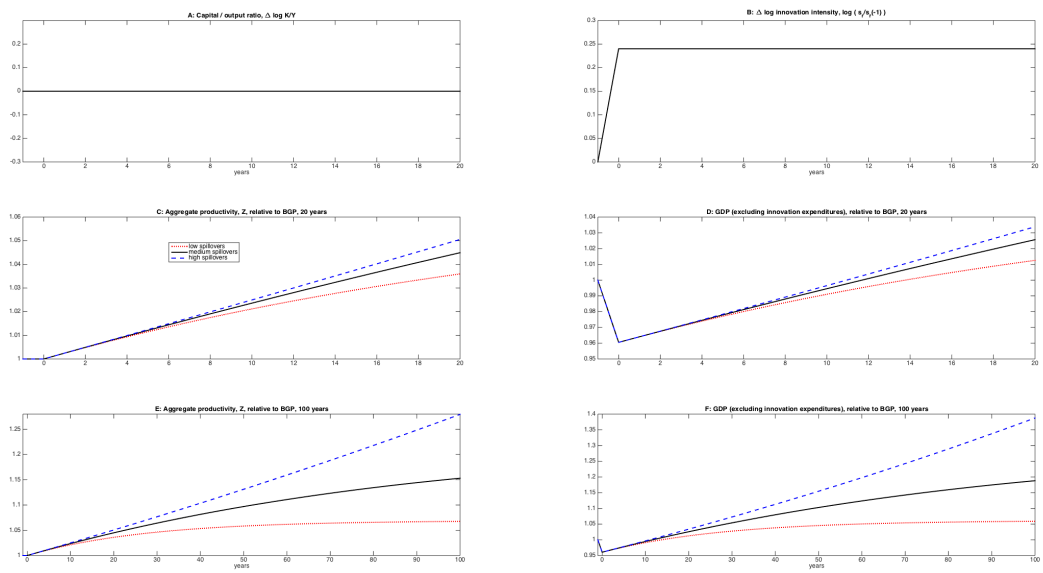

Figure 1: Approximate Transition Path to a Permanent Increase in Innovation Intensity

(ii) Small impact effect Now consider the response of aggregate productivity at the start of the transition. In the middle left and right panels of Figure 1 (panels C and D), we zoom in on the transition of aggregate productivity and GDP for the first 20 years of the transition. Here we see that in the first year of the transition, aggregate productivity rises only a little relative to the initial BGP path (only $\Gamma_0 \times 0.24 = 0.0024$ independent of γ). This implication of the model follows from the bound on the elasticity $\Gamma_0 \leq \bar{g}_z \bar{L}_p$ obtained in Proposition 1.

(iii) Intertemporal knowledge spillovers and medium-term productivity and output dynamics Now consider the response of aggregate productivity over the medium term

(20 years) shown in Panel C of Figure 1. Recall that the elasticity of the cumulative response of aggregate productivity to a permanent change in the innovation intensity of the economy over 20 years is given by $\sum_{k=0}^{19} \Gamma_0$ for the alternative choices of γ. In the first case, as $\gamma \to 1$, we have $\sum_{k=0}^{19} \Gamma_k \to 20 \times \Gamma_0 = .20$ and a response of aggregate productivity at year $t = 20$ of the transition of $\log Z'_{20} - \log \bar{Z}_{20} = .20 \times .24 = 0.049$. In the second case, with $\gamma = 0$, we have $\sum_{k=0}^{19} \Gamma_k = .18$ and $\log Z'_{20} - \log \bar{Z}_{20} = 0.044$. Finally in the third case, with $\gamma = -2$, we have $\sum_{k=0}^{19} \Gamma_k = .147$ and $\log Z'_{20} - \log \bar{Z}_{20} = 0.035$.

Next consider the transition path for GDP exclusive of innovation expenditures (resources available for consumption and physical investment) shown in panel D of Figure 1. In that figure, we see that GDP falls considerably on impact, regains its initial level in 12 years, and is only modestly above its initial level in 20 years. Moreover, we see that the path of GDP in these first 20 years is not particularly sensitive to the choice of the knowledge spillover parameter γ. Recall from Proposition 2 and Corollary 2 that, holding fixed the rental rate of physical capital, the elasticity of GDP at horizon t with respect to a permanent change in the innovation intensity of the economy is given by

$$\frac{\log GDP'_t - \log G\bar{D}P_t}{\log s'_r - \log \bar{s}_r} = \left(\sum_{k=0}^{t-1} \frac{\Gamma_k}{1-\alpha} - \bar{L}_r \right).$$

In our example, this term is equal to $-\bar{L}_r = -.167$ on impact at $t = 0$. With $\gamma \to 1$, this elasticity approaches .138 at $t = 20$, while with $\gamma = -2$, it is .051 at $t = 20$. In our particular policy experiment, since the log change in the innovation intensity of the economy is .24, the log change in GDP excluding innovation expenditures is 0.033 at $t = 20$ with $\gamma \to 1$ and only 0.012 at $t = 20$ with $\gamma = -2$. To convert these results to implications for GDP inclusive of innovation expenditure one must multiply the level of GDP exclusive of these expenditures by $(1 + s_{rt})$.

We see, then, that the model's predictions for the response of aggregate productivity and output to a given permanent change in the innovation intensity of the economy 20 years into the transition are not particularly sensitive to choices of the intertemporal knowledge spillover parameter γ in comparison to the strong dependence of the model's long run predictions for aggregate productivity and output on this parameter.

(iv) Welfare implications and optimal innovation intensity This result that our model's implications for the medium term elasticity of aggregate productivity and GDP with respect to a permanent change in innovation policies is relatively small for a wide range of values of the intertemporal knowledge spillover parameter γ does not imply that the current equilibrium level of innovation expenditures is close to optimal. In fact,

from equation (43), we have that, for a wide range of values of consumers' adjusted discount factor $\tilde{\beta}$ and of the intertemporal knowledge spillover parameters γ, our model implies that the optimal BGP innovation intensity of the economy is higher than the initial level we have assumed. Specifically, using our parameter values above, we have $\mathcal{E}_g^* = \frac{1}{\rho-1}\frac{\exp((\rho-1)\bar{g}_z)-1}{\exp((\rho-1)\bar{g}_z)} = 0.012$. Equation (43) then implies that with $\tilde{\beta} = .99$ (the interest rate is one percentage point higher then the growth rate) and $\gamma = .99$ (close to endogenous growth), the optimal innovation intensity of the economy is $s_r^* = 1.21$, that is, innovation expenditures should exceed expenditure on consumption and physical investment combined by 21%. On the conservative side, with $\tilde{\beta} = .96$ and low intertemporal knowledge spillovers ($\gamma = -2$), $s_r^* = 0.156$, a figure that is not too far from the measures of investments in intangible capital relative to GDP calculated by Corrado et al. (2009). Clearly, our model's implications for the optimal innovation intensity of the economy are highly sensitive to assumptions about consumers' patience and intertemporal elasticity of substitution as summarized by $\tilde{\beta}$ and the level of intertemporal knowledge spillovers γ. Moreover, the model's implications for the optimal innovation intensity of the economy are not particularly sensitive to other parameters.

It is important to note that our model can predict very large welfare gains from moving from an initial BGP with our calibrated innovation intensity of the economy to the socially optimal BGP. The intuition for this implication of our model is evident in Figure 1 panels E and F. There we show the transition for aggregate productivity and output of the final consumption good over the first 100 years. We see that output of consumption falls initially (independent of γ) and then rises after a decade or more, where the magnitude of the rise is quite sensitive to γ. As a consequence of these differences in the response of output of the final consumption good in the long run, we find that the welfare implications of an increase in innovation subsidies are quite sensitive to the degree of intertemporal knowledge spillovers that we assume. In a full numerical solution of the model presented in the appendix (that also takes into account the transition dynamics of physical capital) we find in the low spillover case that the equivalent variation in consumption is 3.9%, in the medium spillover case it is 13.9%, and in the high spillover case it is 41.1%. Thus, with large intertemporal knowledge spillovers, we find very large welfare gains from the innovation subsidies in this experiment. The differences in the welfare implications of innovation policy changes across the high and low intertemporal knowledge spillover economies arise as a result of the growing differences in the paths of consumption past the medium term horizon of 20 years in combination with our choice of $\tilde{\beta} = .99$ for the consumer's adjusted-discount factor.

Sensitivity to alternative parameter choices Our characterization of our model's dynamics using simple first order approximations is useful for highlighting which features of the model are important in determining its quantitative implications and for understanding quantitatively the sensitivity of these implications with respect to these model features. For example, consider the impact of calibrating the model to a higher level of productivity growth $\bar{g}_z$ on the initial BGP. If we double this initial level of productivity growth, holding other parameters fixed (except for g_{A_r}, which we modify to be consistent with this higher growth rate), then the upper bound on the elasticity of aggregate productivity on impact Γ_0 also doubles. The change in the exact value of Γ_0 depends on the value of ρ: for ρ close to 1, Γ_0 also doubles; as ρ approaches infinity, the value of Γ_0 becomes insensitive to $\bar{g}_z$. With $\rho = 4$, doubling $\bar{g}_z$ from .0125 to .025 raises Γ_0 from 0.01 to .02. In the case of endogenous growth, $\gamma = 1$, the 20-year elasticity of aggregate productivity $\sum_{k=0}^{19} \Gamma_k$ increases from .2 to .4. In the semi-endogenous growth case, $\gamma < 1$, the terms Γ_k decay more quickly, so, in the case with low intertemporal knowledge spillovers $\gamma = -2$ that we considered, this medium term elasticity rises from .147 to .215 Hence in this case, the elasticities of productivity implied by the model are larger and the medium term implications of the model are more sensitive to changes in the knowledge spillover parameter γ.

Likewise, consider the sensitivity of our results to the calibration of the innovation intensity of the economy $\bar{s}_r$ on the initial BGP. If we calibrate the model to a lower initial value of $\bar{s}_r$ on the initial GDP and increase the innovation intensity of the economy by 3 percent of GDP (i.e. $s' - \bar{s}_r = 0.03$) then, mechanically, our model implies that this policy experiment results in a larger change in the *log* of the innovation intensity of the economy. Thus, keeping the model elasticities Γ_k unchanged, the magnitude of the response of aggregate productivity to this policy experiment will be larger. For example, if we had assumed an initial innovation intensity of the economy of 5 percent in line with the new measures of intangible investment relative to Private Business Output in the NIPA in recent years, then our policy experiment would have increased the log of the innovation intensity of the economy by $\log .08 - \log .05 = .47$ rather than .24 so that the magnitude of the aggregate responses, up to a first-order approximation, would roughly double.

Example 2: social depreciation of innovation expenditures As we noted above, the Expanding Varieties model that we considered in example 2 builds in the implicit assumption that there is social depreciation of innovation expenditures and thus implies a larger impact elasticity $\mathcal{E}^{EV}$ than in the Quality Ladders model. As we discussed above, the specific social depreciation rate of innovation expenditures implied by the model is

pinned down in data by the choice of the parameter, ρ, the baseline growth rate of aggregate productivity, $\bar{g}_z$, and the calibrated share of employment in incumbent firms on the BGP, $(1-\delta_f)\exp((\rho-1)\Delta_z)/\exp((\rho-1)\bar{g}_z)$.

Consider the implications of this model if it is calibrated similarly to the Quality Ladders model in example 1 in that we set $\rho = 4$ and calibrate the initial BGP to have $\bar{g}_z = 0.0125$ and $\bar{L}_p = 0.83$. If we assume that the share of employment in incumbent firms (or incumbent products) is 90%, we get that aggregate productivity would decline by roughly 2.25% if the innovation rate was zero (i.e. the social rate of depreciation of innovation expenditures is $G^0 = -0.0225 = \frac{1}{\rho-1}[\log(0.9) + (\rho-1)\bar{g}_z]$). In this case, we have an impact elasticity $\mathcal{E}^{EV}_{g0} = 0.033$ which is roughly three times larger than the elasticity of $\mathcal{E}^{QL}_{g0} = .0122$ that we found in the Quality Ladders model in which there is no social depreciation. With these assumptions, together with our initial calibration of $\bar{L}_p$, we get $\Gamma_0 = .028$ rather than the value of $\Gamma_0 = .01$ that we obtained under the assumption of no social depreciation of innovation. This implies that with high intertemporal knowledge spillovers ($\gamma \to 1$) the medium term elasticity of aggregate productivity with respect to a permanent change in the innovation intensity of the economy is $\sum_{k=0}^{19}\Gamma_k = .56$ rather than 0.2 with no social depreciation of innovation. With low intertemporal knowledge spillovers ($\gamma = -2$), this elasticity is $\sum_{k=0}^{19}\Gamma_k = .24$ rather than .147 with no social depreciation of innovation. On the other hand, $\sum_{k=0}^{\infty}\Gamma_k = \frac{\bar{L}_p}{1-\gamma}$, which determines the long-term change (across BGPs) in aggregates, is independent of the social rate of depreciation.

These results imply that in an economy with a moderate initial productivity growth rate and with a modest social depreciation of innovation, the medium term elasticity of aggregate productivity with respect to a permanent change in the innovation intensity of the economy is both significantly larger than it is with no social depreciation of innovation and much more sensitive to the magnitude of intertemporal knowledge spillovers. In particular, with significant social depreciation of innovation and high intertemporal knowledge spillovers, we have from Corollary 2 that measured GDP exclusive of innovation expenditures rises relative to its initial BGP path after only 5 years of transition following a permanent increase in the innovation intensity of the economy instead of after more than 10 years as shown in panel D of Figure 1.

Example 3: adding innovation by incumbents We now turn to the model example 3, the Klette Kortum model with innovation by both incumbents and entrants. Proposition 3 states that, starting from a conditionally efficient BGP and a given allocation of labor $\bar{L}_p$, this model implies a smaller elasticity of aggregate productivity to the innovation intensity of the economy along the transition path than the Quality Ladders version of

the model in which innovation is only carried out by entrants (model example 1). The impact elasticity in this model is lower than in the Quality Ladders model by a factor $\bar{Y}_r/(\bar{Y}_r - Y_r^0) < 1$. In the model calibration described in the Appendix, the conditionally efficient allocation implies $\bar{Y}_r/(\bar{Y}_r - Y_r^0) = .966$, very close to its maximum value of one. Therefore, the aggregate dynamics in response to a given path of the innovation intensity of the economy is almost indistinguishable from that in the Quality Ladders model.

Lack of conditional efficiency at the initial baseline allocation The models in examples 3 and 4 allow for the possibility that the initial allocation of research expenditure across entering and incumbent firms is not conditionally efficient. As we discussed in Section 4 ("Lack of conditional efficiency and the impact elasticity") if the initial allocation of the research good between entrants and incumbents on the initial baseline BGP is not conditionally efficient, it is possible to attain the same BGP growth rate with a lower aggregate output of the research good and increase consumption to a new permanently higher level by increasing labor in the production of intermediate goods. The quantitative magnitude of such gains are difficult to restrict without detailed data on the specifics of the innovation technologies available to different types of firms. To illustrate this point, consider the simplest variation on the baseline allocation of the research good across entering and incumbent firms that we assumed in our model in example 3. Specifically define a misallocation "wedge" τ different from 1 such that $\bar{y}_r(0)/d'(\bar{y}_r(1)) = \tau$ in the initial BGP. Then, in a calibration of the innovation technology for incumbents $d(y_r(1))$ described in the Appendix, the increase in production labor (and hence consumption once physical capital fully adjusts) that can be accomplished while maintaining the same growth rate of aggregate productivity by moving to the conditionally efficient allocation varies significantly with τ. For example, the permanent percentage increase in production labor is 0% if $\tau = 1$ (i.e in this case the initial allocation is conditionally efficient), 0.1% if $\tau = 1.1$, 1.4% if $\tau = 1.3$, and 6.8% if $\tau = 1.5$. We see that if the distribution of innovative investments across firms is sufficiently far from conditional efficiency, then it is feasible to reallocate a substantial amount of labor from research to current production without reducing the growth rate of aggregate productivity.

The challenge to developing policies that improve welfare primarily by reallocating innovative investments across firms is to identify which firms should be doing relatively more innovation spending and by how much should these firms increase their investments in innovation. Answering these questions requires detailed knowledge of the technology linking innovative investments by different firms to aggregate productivity growth. There are several papers in the literature that have taken up this challenge by

estimating full models of firm dynamics using micro data and structural models based on the Klette-Kortum model framework. We now examine the results from one of those studies in the light of the simple approximations we present above. We see the calculations we undertake below as a useful metric on how far a specific estimated model gives different results than the baseline models considered in this paper.

Example 4: Gains from reallocation in the Lentz-Mortensen model: Our model in example 4 is from the paper by Lentz and Mortensen (2014). In this paper, these authors consider an endogenous growth version of this model, so that $\gamma = 1$. They estimate the other parameters of the model of example 4 with incumbent firms with two types of innovation technologies using firm level data. They then compute the change in growth rate of aggregate productivity that can be achieved by moving from what they estimate to be the initial BGP to the BGP in which firm entry and innovative investment decisions are optimally chosen. In table 2 on page 15 of this paper (the draft of December 11, 2014), in the first two columns of that table, the authors report their results. They find that the growth rate of aggregate productivity on the socially optimal BGP is twice as high as that on the initial BGP: 2.8% rather than 1.4%. This is associated with a substantial reallocation of innovation expenditure across firms, a large change in the distribution of incumbent firms by type in the long run, and a large increase in aggregate production of the research good. They find that the welfare gains from moving from the initial BGP to this socially optimal BGP are very high: on the order of 20% of aggregate consumption. How do these estimated results compare to the simple bounds that we found relating the change in aggregate production of the research good to the change in the BGP growth rate of aggregate productivity in endogenous growth versions of our model? Specifically, how does the change in the growth rate across BGP's ($.014 = .028 - .014$) compare to the bound for models that satisfy assumptions 1-3 of the initial baseline growth rate of aggregate productivity $\bar{g}_z$ times the log change in aggregate production of the research good across BGP's. (Note that the authors assume that $\rho = 1$ so that the bounds we derive under assumptions 1-3 are tight.)

As discussed above, the initial baseline allocation in this Lentz-Mortensen model is not conditionally efficient. Moreover, the distribution of incumbent firms by type enters into the function G describing the relationship between innovation spending by firms and the growth rate of aggregate productivity. Thus, this model violates both assumptions 1 and 3 above. These deviations of the specification of this model from our assumptions may allow this model to have a larger response of the aggregate growth rate to any given change in the aggregate production of the research good as the economy moves from the

initial BGP to the socially optimal BGP. At the same time, these authors assume that both incumbent and entering firms have strictly concave technologies $d(y_r)$ for producing innovations, and thus the function G in this model is more concave than it is in the model of example 1 which has a linear entry margin. This second assumption works to require a larger change in aggregate production of the research good to accomplish a given change in the growth rate of aggregate productivity than would be the case in models with a linear margin for entering firms. Thus, in their estimated model, these authors can find a change in the growth rate of aggregate productivity across BGP's in the long run that is either larger or smaller than the initial baseline growth rate of aggregate productivity $\bar{g}_z$ times the log change in aggregate production of the research good across BGP's. What these authors find, as reported in the first two columns of Table 2 in their paper, is that these two effects roughly cancel out. The log of aggregate production of the research good (research labor) increases by $1.12 = \log(4.856) - \log(1.582)$ and thus their estimated long run change in aggregate productivity growth of $0.014 = 0.028 - 0.014$ is only 0.0125 times the log change in aggregate production of the research good as the economy moves from its initial BGP to the socially optimal BGP. This ratio is slightly smaller than their baseline value of $\bar{g}_z = 0.014$. This ratio is slightly smaller than their baseline calibration of aggregate productivity growth of $\bar{g}_z = 0.014$, which also corresponds to the elasticity of the productivity growth rate with respect to changes in the research we obtained analytically in the baseline model of example 1.

6 Conclusion

In this paper, we have derived a simple first-order approximation to the transition dynamics of aggregate productivity and GDP in response to a policy-induced change in the innovation intensity of the economy implied by a model that nests a widely-used class of models of growth through firms' investments in innovation that have been developed over the past 25 years. Our results highlight the key features/parameters of these models that drive their quantitative predictions for the aggregate implications of innovation policies.

Our first result analyzed the immediate response of aggregate productivity growth to an increase in aggregate production of the research good — a response that we termed the impact elasticity of aggregate productivity growth with respect to a change in aggregate production of the research good. We showed that if the aggregate technology relating firms' investments in innovation to aggregate productivity growth satisfies a concavity assumption and if the initial baseline allocation at which the policy change is considered

is conditionally efficient, then this impact elasticity is bounded by the gap between the growth rate of aggregate productivity in the baseline allocation and the social depreciation of innovation expenditures, defined as the growth rate of aggregate productivity that would obtain if all firms in the economy invested nothing in innovation. Hence, specifications of our model that assume no social depreciation of innovation expenditures and which are calibrated to a low initial baseline growth rate of aggregate productivity characteristic of advanced economies necessarily imply a low impact elasticity of aggregate productivity growth with respect to a change in aggregate production of the research good if the baseline allocation is conditionally efficient. In contrast, in specifications of our model which for social depreciation, then this impact elasticity has the potential to be substantially larger.

Our results bounding the impact elasticity of aggregate productivity growth with respect to a change in aggregate production of the research good relied on the assumption that the technology G linking innovative investments by firms to aggregate productivity growth satisfied a certain concavity assumption. This assumption rules out contemporaneous increasing returns in the aggregate to firms' joint innovative investments. We do not take a stand on whether this assumption is satisfied in reality. We highlight this assumption simply because it characterizes the aggregate technology G assumed in a widely used class of models.

We illustrated that Neo-Schumpeterian models built on the Quality Ladders framework differ from Expanding Varieties model in that the latter model specification implicitly assumes that there is social depreciation of innovation expenditures. We do not intend this illustration to stand as a definitive theory of the social depreciation of innovation expenditures. We speculate that social depreciation of innovation expenditures is likely derived from the fact that productive knowledge in firms is actually embodied in individuals who are familiar both with the knowledge gained through innovation and the procedures for training new workers in that knowledge. We imagine that a full theory of the social depreciation of innovation expenditures would have to account for the dynamics of social learning that goes on within firms in an environment in which the work force within firms is constantly turning over and workers themselves have a life cycle. We leave such a study to future research.

We also examined the implications of considering specifications of our model that include innovative investments by both incumbent and entering firms for our model's predictions for the impact elasticity of aggregate productivity growth with respect to changes in the aggregate production of the research good. We showed this elasticity is smaller than in equivalent models with innovative investments only by entering firms if

incumbent firms are assumed to have a lower average cost of contributing to aggregate productivity growth than entrants. Specific quantitative statements regarding the magnitude of this reduction of the impact elasticity would require detailed knowledge of the technology linking investments by firms to aggregate productivity growth.

We next examined the dynamics of aggregate productivity and GDP in response to persistent changes in the innovation intensity of the economy induced by persistent changes in innovation policies. Under the assumption made in the most tractable specifications of our model that the technology linking firms' innovative investments to aggregate productivity growth does not depend on the distribution of incumbent firms in the economy by type, we showed that these transition dynamics are characterized entirely by the impact elasticity of aggregate productivity growth with respect to a change in aggregate production of the research good, the baseline share of labor employed in production $\bar{L}_p$, and the parameter γ governing the magnitude of intertemporal knowledge spillovers in the production of the research good. We showed that the implications of our model for the elasticity of aggregate productivity and output over the long term to permanent changes in the innovation intensity of the economy is determined largely by the assumptions made about the intertemporal spillovers of knowledge in the research sector. We showed that in specifications of our model with a low impact elasticity of aggregate productivity with respect to changes in aggregate production of the research good (due to assumptions of no social depreciation of innovation expenditures and/or a low initial growth rate of aggregate productivity), alternative assumptions about the extent of these intertemporal knowledge spillovers (measured in terms of our model's long term implications for aggregate productivity and output) have only a small impact on the model's implications for elasticities of aggregate productivity over the medium term. These results imply that it may be difficult to identify the extent of intertemporal knowledge spillovers in the time series data that are available to us.

We also used these results about model dynamics to show that the implications of our model for the socially optimal level of the innovation intensity of the economy are driven primarily by the assumed intertemporal knowledge spillovers in the production of the research good and the assumed patience of consumers relative to the growth rate of the economy. If intertemporal knowledge spillovers are large, then a permanent increase in the innovation intensity of the economy can generate very large increases in aggregate productivity and output in the long run. If consumers are patient, they view this long run payoff as having a large benefit relative to the short run cost of lowering current consumption to allow for increased investment in innovation. Given our uncertainty regarding the magnitude of intertemporal knowledge spillovers in the production of re-

search, it seems difficult to use our model to make even rough quantitative statements about the optimal innovation intensity of an economy. In terms of practical applications, this uncertainty about the magnitude of intertemporal knowledge spillovers corresponds to uncertainty about the counterfactual movements in the relative price of the research good (the deflator of innovation expenditures) that would result from a policy induced permanent increase in the innovation intensity of the economy.

Under our baseline assumptions, we have abstracted from the productivity and welfare gains that might be achieved by reallocating a given level of investment in innovation across heterogeneous firms. As we saw in our consideration of misallocation of research expenditure across firms in the context of the simple Klette-Kortum model of example 3, the welfare gains from such a reallocation can potentially be large. Under alternative model specifications in which there are first-order gains from such a reallocation, one may wish to consider a whole range of policies aimed at reallocating innovation expenditures across firms. The research challenge here is to find reliable metrics for evaluating which firms should be doing more investment in innovation and which should be doing less and what the nature of their different technologies for contributing to the aggregate productivity growth rate might be.

References

Acemoglu, Daron, "Introduction to Modern Economic Growth," 2009.

_ **, Ufuk Akcigit, Nicholas Bloom, and William R. Kerr**, "Innovation, Reallocation and Growth," NBER Working Papers 18993, National Bureau of Economic Research, Inc April 2013.

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, March 1992, *60* (2), 323–51.

_ **, Ufuk Akcigit, and Peter Howitt**, "What Do We Learn From Schumpeterian Growth Theory?," February 2013, (298).

Aizcorbe, Ana, Carol Moylan, and Carol Robbins, "Toward Better Measurement of Innovation and Intangibles," BEA Briefing, Bureau of Economic Analysis U.S. Department of Commerce 2009.

Akcigit, Ufuk and William R. Kerr, "Growth Through Heterogeneous Innovations," November 2010, (16443).

_ **, Douglas Hanley, and Nicolas Serrano-Velarde**, "Back to Basics: Basic Research Spillovers, Innovation Policy and Growth," NBER Working Papers 19473, National Bureau of Economic Research, Inc September 2013.

Atkeson, Andrew and Ariel Tomas Burstein, "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 06 2010, *118* (3), 433–484.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 07 2013, *81* (4), 1347–1393.

Buera, Francisco and Roberto Fattal-Jaef, "The Dynamics of Development: Entrepreneurship, Innovation, and Reallocation," Working Paper, World Bank 2014.

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

Corrado, Carol A. and Charles R. Hulten, "Innovation Accounting," in "Measuring Economic Sustainability and Progress" NBER Chapters, National Bureau of Economic Research, Inc, December 2013.

Corrado, Carol, Charles Hulten, and Daniel Sichel, "Measuring Capital and Technology: An Expanded Framework," in "Measuring Capital in the New Economy" NBER Chapters, National Bureau of Economic Research, Inc, December 2005, pp. 11–46.

_ , _ , and _ , "Intangible Capital And U.S. Economic Growth," *Review of Income and Wealth*, 09 2009, *55* (3), 661–685.

_ , Peter Goodridge, and Jonathan Haskel, "Constructing a Price Deflator for R&D: Calculating the Price of Knowledge Investments as a Residual," Economics Program Working Papers 11-03, The Conference Board, Economics Program October 2011.

Garcia-Macia, Daniel, Pete Klenow, and Chang-Tai Hsieh, "How Destructive is Innovation?," January 2015.

Griliches, Zvi, "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, Spring 1979, *10* (1), 92–116.

_ , "The Search for R&D Spillovers," in "R&D and Productivity: The Econometric Evidence" NBER Chapters, National Bureau of Economic Research, Inc, December 1998, pp. 251–268.

_ , ed., *R&D, Patents and Productivity* number 9780226308845. In 'National Bureau of Economic Research Books.', University of Chicago Press, 1987.

Grossman, Gene and Elhanan Helpman, *Innovation and Growth in the Global Economy*, MIT Press, 1991.

Grossman, Gene M and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, January 1991, *58* (1), 43–61.

Jaimovich, Nir and Sergio Rebelo, "Non-linear Effects of Taxation on Growth," NBER Working Papers 18473, National Bureau of Economic Research, Inc October 2012.

Jones, Charles I., "Sources of U.S. Economic Growth in a World of Ideas," *American Economic Review*, March 2002, *92* (1), 220–239.

_ , "Growth and Ideas," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1 of *Handbook of Economic Growth*, Elsevier, 2005, chapter 16, pp. 1063–1111.

_ and John C. Williams, "Measuring The Social Return To R&D," *The Quarterly Journal of Economics*, November 1998, *113* (4), 1119–1135.

Jorgenson, Dale, Mun Ho, and Jon Samuels, "Long-Term Estimates of U.S. Productivity and Growth.," Working Paper 2014.

Kendrick, John, "Total Capital and Economic Growth," *Atlantic Economic Journal*, 1994, *22* (1), 1–18.

Klette, Tor Jakob and Samuel Kortum, "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, October 2004, *112* (5), 986–1018.

Kortum, Samuel S., "Research, Patenting, and Technological Change," *Econometrica*, November 1997, *65* (6), 1389–1420.

Lentz, Rasmus and Dale T. Mortensen, "An Empirical Model of Growth Through Product Innovation," *Econometrica*, November 2008, *76* (6), 1317–1373.

__ **and** __ , "Optimal Growth through Product Innovation," 2014.

Li, Wendy, "Technological Innovation, Resource Allocation, and Growth," R and D Satellite Accounts Paper, Bureau of Economic Analysis U.S. Department of Commerce 2012.

Luttmer, Erzo G. J., "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 08 2007, *122* (3), 1103–1144.

McGrattan, Ellen R. and Edward C. Prescott, "Taxes, Regulations, and the Value of U.S. and U.K. Corporations," *Review of Economic Studies*, 2005, *72* (3), 767–796.

__ **and** __ , "On financing retirement with an aging population," Staff Report 472, Federal Reserve Bank of Minneapolis 2012.

Peters, Michael, "Heterogeneous Mark-Ups, Growth and Endogenous Misallocation," 2013.

Rivera-Batiz, Luis A and Paul M Romer, "Economic Integration and Endogenous Growth," *The Quarterly Journal of Economics*, May 1991, *106* (2), 531–55.

A Equilibrium of baseline Klette-Kortum model (example 3)

The numerical examples in the body of the paper are based on a calibration of our baseline Klette-Kortum model (example 3). In the body of the paper we presented a partial characterization of the equilibrium conditions that we use to derive our analytic results. In this appendix we complete the description of the equilibrium of this model and its solution both in the BGP and out-of BGP. We also provide a proposition relating changes in fiscal expenditures and changes in the innovation intensity of the economy.

Characterizing and solving equilibrium

Variable profits that an incumbent firm earns in period t from production of a product with productivity index z are given by $[p_t(z)y_t(z) - W_t l_t(z) - R_{kt}k_t(z)]$. The incumbent firm that owns this product chooses price and quantity, $p_t(z)$ and $y_t(z)$, to maximize these variable profits subject to the demand for its product and the production function (3). We assume that the gross markup μ charged by the incumbent producer of each product is the minimum of the monopoly markup, $\rho/(\rho-1)$, and the technology gap between the leader and any potential second most productive producer of the good, $\exp(\Delta_l)$. That is $\mu = \min\left\{\frac{\rho}{\rho-1}, \exp(\Delta_l)\right\}$. Variable profits from production can then be written as $\Pi_t \exp(z)^{\rho-1}$, with the constant in variable profits Π_t defined by

$$\Pi_t = \kappa_0 \left(R_{kt}^{\alpha} W_t^{1-\alpha}\right)^{1-\rho} Y_t, \tag{44}$$

with $\kappa_0 = \mu^{-\rho}(\mu-1)\left[\alpha^{\alpha}(1-\alpha)^{1-\alpha}\right]^{\rho-1}$.

Intermediate goods producing firms are offered two types of innovation subsidies. First, incumbent firms receive a subsidy to innovation expenditure, which we denote by τ_g. Second, entering firms receive a subsidy to innovation expenditure, which we denote by τ_e. We refer to policies in which $\tau_g = \tau_e$ as *uniform innovation subsidies.*

The expected discounted stream of profits associated with selling a product with productivity index z is given by the solution to the following Bellman equation which takes into account the probability that the firm loses its ownership of this product to another innovating firm

$$V_t(z) = \Pi_t \exp(z)^{\rho-1} + \frac{(1-\delta_t)}{1+\bar{r}_t} V_{t+1}(z),$$

where δ_t denotes the total measure of products innovated on at time t (which every firm takes as given), $\bar{r}_t$ denotes the interest rate denominated in the final consumption good, which with log preferences is given by $1+\bar{r}_t = \beta^{-1}C_{t+1}/C_t$. Integrating this expression across z and using the

definition of the aggregate Z_t, we have

$$V_t Z_t^{\rho-1} = \Pi_t Z_t^{\rho-1} + \frac{(1-\delta_t)}{1+\bar{r}_t} V_{t+1} Z_t^{\rho-1}, \tag{45}$$

where $V_t = V_t(0)$.

The second component of the value to a firm of owning a product is given by the expected present value of dividends the incumbent firm expects to earn on new products it gains through innovation minus the cost of that innovation. Given that this value depends on the number of products owned by the firm, n, and not on its productivities, we denote it by $U_t(n)$. Conjecturing that $y_{rt}(n) = y_{rt}(1)$ (we verify this below), the value U_t is determined by the Bellman equation

$$\begin{aligned} U_t &= \max_{\tilde{q}} -(1-\tau_g) y_{rt}(1) P_{rt} + \frac{1}{1+\bar{r}_t} \sigma d(y_{rt}(1)) V_{t+1} Z_t^{\rho-1} \exp(\Delta_z)^{\rho-1} \\ &\quad + \frac{1}{1+\bar{r}_t} (\sigma d(y_{rt}(1)) + 1 - \delta_t) U_{t+1}. \end{aligned} \tag{46}$$

The first term on the right side indicates the firm's investment in innovation. The second term indicates the discounted present value of variable profits the firm expects to gain from the innovations that result from this investment. The third term denotes the expected value of the firm's innovative capacity from next period on, taking into account both the gain in products it expects to obtain from its innovative effort (i.e., a firm with n products expects to gain $\sigma d(y_{rt}(1)) n$ products) and the loss of products it expects as a result of innovative effort from other firms (i.e., a firm with n products expects to lose δn products). Note that in equilibrium, if the incumbents' innovation technology is used, we must have $U_t \geq 0$. Otherwise, incumbents would choose not to use their innovation technology at all. Hence, given sufficiently low values of the function $d(.)$ such that incumbents choose not to innovate, this model nests the standard quality ladders model.

The first-order condition from equation (46) for optimal innovative investment per product by incumbent firms, $y_{rt}(1)$ (taking δ_t is given) is given by

$$(1-\tau_g) P_{rt} = \left(\frac{\sigma d'(y_{rt}(1))}{1+\bar{r}_t} \right) \left(V_{t+1} Z_t^{\rho-1} \exp(\Delta_z)^{\rho-1} + U_{t+1} \right). \tag{47}$$

Since none of the terms depend on the incumbent firm's number of products n or the productivities with which the firm can produce those products, we verify our conjecture that that per-product innovation spending, $y_{rt}(1)$, is the same for all products and firms.[24] The total measure

[24]One can show that the property that all incumbent firms choose the same $y_{rt}(1)$ extends in an alternative specification of our model in which each product that is innovated on draws a random markup and innovation step size that is independent of the identity of the innovator. Under this extension, consistently with the data as emphasized in Klette and Kortum (2004), the model generates persistent variation in labor revenue productivity across firms and in research intensity (defined as innovation expenditures relative to revenues). In particular firms that have, on average, products with higher markups have higher measured labor productivity. Moreover, since all firms choose the same level of innovative investment per product, measured labor productivity and research intensity are not correlated with firm growth in terms of its

of products innovated on is $\delta_t = \sigma\left(y_{rt}(1) + N_t(0)\right)$.

The total value of an incumbent firm with n products with frontier technologies $z_1, ..., z_n$ is the sum of the values over its current products, $\sum_{i=1}^{n} V_t \exp(z_i)^{\rho-1}$ plus the value of its innovative capacity, $U_t n$. The free entry condition for new firms is given by

$$(1-\tau_e) P_{rt}\bar{y}(0) \geq \left(\frac{\sigma}{1+\bar{r}_t}\right)\left(V_{t+1} Z_t^{\rho-1} \exp(\Delta_z)^{\rho-1} + U_{t+1}\right), \tag{48}$$

with this condition holding with equality if the measure of entering firms, $N_t(0)$, is greater than zero. If the equilibrium has entering firms, then the zero-profit condition for entry (48) and equation (47) imply that incumbents' innovation per product is determined from

$$\frac{(1-\tau_g)}{d'(y_{rt}(1))} = (1-\tau_e)\,\bar{y}_r(0). \tag{49}$$

This result implies that in any equilibrium with positive entry and uniform innovation subsidies ($\tau_g = \tau_e$), $d'(y_{rt}(1))\,\bar{y}_r(0) = 1$, which coincides with the condition for conditionally efficiency.

Rescaled Bellman equations: It is useful to present rescaled Bellman equations describing the value of firms. Defining $v_t = \frac{V_t Z_t^{\rho-1}}{P_{rt}}$, $u_t = \frac{U_t}{P_{rt}}$, we have (without imposing BGP),

$$v_t = \frac{\Pi_t Z_t^{\rho-1}}{P_{rt}} + \frac{(1-\delta_t)}{1+r_t}\frac{v_{t+1}}{\exp((\rho-1)\,g_{zt})}, \tag{50}$$

$$\begin{aligned} u_t &= -(1-\tau_g)\,y_{rt}(1) + \frac{1}{1+r_t}\frac{1}{\exp((\rho-1)\,g_{zt})}\sigma d(y_{rt}(1))\,v_{t+1}\exp(\Delta_z)^{\rho-1} \\ &\quad + \frac{1}{1+r_t}\left(\sigma d(y_{rt}(1)) + 1 - \delta_t\right) u_{t+1}. \end{aligned} \tag{51}$$

and the free entry condition is

$$(1-\tau_e)\,\bar{y}_r(0) \geq \left(\frac{\sigma}{1+r_t}\right)\left(\frac{v_{t+1}}{\exp((\rho-1)\,g_{zt})}\exp(\Delta_z)^{\rho-1} + u_{t+1}\right), \tag{52}$$

where $1 + r_t = (1+\bar{r}_t)\,\frac{P_{rt}}{P_{rt+1}}$.

Solving BGP: In a BGP, the constant in rescaled profits $\Pi Z^{\rho-1}/P_r$, and values $VZ^{\rho-1}/P_r$, U/P_r are constant over time (as well as r, g_z and δ as discussed above).

In the semi-endogenous growth case ($\gamma < 1$) the rate at which innovations occur, δ, is pinned down from equation (12). Innovation by incumbent firms $y_r(1)$ is constant and, with positive firm entry, is determined as the solution to (49), and entry is given by $N(0) = \frac{\delta}{\sigma} - y_r(1)$. In

number of products. This extension of the model does not change substantially our main results.

what follows we focus on parameter values such that on the BGP, new firms enter ($N(0) > 0$ when calculated as described above) and incumbents choose to use their innovation technology. Incumbent firms find it optimal to use their innovation technology in a BGP with firm entry ($u > 0$) if and only if $\left(1-\tau_g\right)\frac{y_r(1)}{d(y_r(1))} \leq \left(1-\tau_e\right)\bar{y}_r(0)$. By equation (49), this condition is equivalent to $d'\left(y_r(1)\right) < \frac{d(y_r(1))}{y_r(1)}$, which is satisfied if $d(.)$ is concave and $d(.) = 0$.

The constant $\Pi Z/P_r$ is pinned down by the free entry condition as follows. Manipulating equations (50) and (51) and imposing BGP with positive entry, we can write the free-entry condition (52) as

$$\left(1-\tau_e\right)\bar{y}_r(0)\,N(0) = \frac{1}{1+\frac{r}{\delta}\left(d\left(y_r(1)\right)+1\right)}\left[\xi_g\frac{\Pi Z^{\rho-1}}{P_r} - \left(1-\tau_g\right)y_r(1)\right], \tag{53}$$

$$\text{where } \xi_g = \frac{r+\delta}{r\exp\left(\left(\rho-1\right)g_z\right)/\exp\left(\Delta_z\right)^{\rho-1}+\delta} \geq 1$$

and $r = \frac{1+\bar{r}}{\exp(\bar{g}_y)} - 1 = \tilde{\beta}^{-1} - 1$ in the BGP (since in the BGP, P_{rt} falls at a rate $\bar{g}_y$). Given the value of $\Pi Z^{\rho-1}/P_r$ that solves this equation, the allocation of labor L_p/L_r is determined as a function of parameters using

$$\frac{L_p}{L_r} = \left(\frac{1-\alpha}{\mu-1}\right)\frac{\Pi Z^{\rho-1}}{P_r}\frac{1}{y_r(1)+\bar{y}_r(0)\,N(0)}. \tag{54}$$

The level of aggregate productivity Z, for a given current value of scientific knowledge A_r, is determined using equation (8). When $\gamma = 1$, one can use the same procedure but instead of solving for Z (the BGP level of Z, given A_r, is not pinned down), one must solve for the growth rate g_z (and the corresponding values of δ, $y_r(1)$ and $N(0)$). The innovation intensity of the economy, s_r, is calculated as a function of parameters using expression (15). Finally, we solve for aggregate output, Y, using (4), the stock of physical capital, K_t, using the factor shares of physical capital and production-labor, and consumption, C_t, using (1).

Solving transition dynamics In Sections 4 and 5 section we approximated the aggregate transition taking as given the transition path for the innovation intensity of the economy and the ratio of physical capital to output. To solve for the path of these two variables for a given change in policies or other parameters, we solve the model numerically. Specifically, we solve for the path of Z_t, K_t, v_t, and u_t using the four following Euler equations: (50), (51), (52) and $R_{kt} = d_k + \bar{r}_t$ where $\bar{r}_t = \beta^{-1}\left(\frac{C_{t+1}}{C_t}\right)^{\xi} - 1$. Recall that $y_{rt}(1)$ is solved for using equation (49) assuming that there is positive firm entry (which must be checked). Given a path of Z_t and K_t we can solve for all other equilibrium outcomes using static equations. We solve for the 4 Euler equations using either standard linearization methods or a shooting algorithm, and we obtain very similar results.

Fiscal expenditures and innovation intensity

The following proposition derives the elasticity across BGPs of the innovation intensity of the economy and fiscal expenditures on innovation policies with respect to a uniform change in innovation policies.

Proposition 4. *Consider our baseline Klette-Kortum economy (like example 3) on a BGP with semi-endogenous growth and positive firm-entry. Suppose that innovation policies change permanently from $\bar{\tau}_g = \bar{\tau}_e$ to $\tau_g' = \tau_e'$. Then, across the old and new BGP the innovation intensity of the economy changes from $\bar{s}_r$ to s_r', and fiscal expenditures relative to GDP change from $\bar{E}/G\bar{D}P$ to E'/GDP', with these changes given by*

$$\log s_r' - \log \bar{s}_r = \log(1 - \bar{\tau}_g) - \log(1 - \tau_g')$$

and

$$\frac{E'}{GDP'} - \frac{\bar{E}}{G\bar{D}P} = s_r' - \bar{s}_r.$$

Proof. Under the assumption of semi-endogenous growth ($\gamma < 1$), δ, g_z, r, and ξ_g are constant between BGPs. Under uniform innovation policies and firm-entry, $y_r(1)$, $N(0)$ and Y_r are constant between BGPs. By equation (53), $\frac{\Pi Z^{\rho-1}}{P_r}\frac{1}{1-\tau_g}$ must be constant between BGPs. Using $\Pi_t Z_t^{\rho-1} = \frac{\mu-1}{\mu} GDP_t$ and the fact that Y_r is constant between BGPs, $\frac{1}{s_r\left(1-\tau_g\right)}\frac{(\mu-1)}{\mu}$ must be unchanged between BGPs, which immediately implies the first result of the proposition. The second result follows from the fact that with uniform innovation policies we have $\frac{E}{GDP} = \tau_g s_r$, combined with the previous result that $s_r\left(1-\tau_g\right)$ is constant between BGPs. □

It is straightforward to derive the same Proposition 4 in the expanding varieties version of our model (either with or without incumbents' innovation). Proposition 4 implies that in the long-run, our policy experiment will result in a change in the innovation intensity of the economy determined only by the change in the innovation subsidy rate independent of the other parameters of the model. At short and medium horizons, however, this policy will result in a change in the path of the innovation intensity of the economy from $\{\bar{s}_{rt}\}_{t=0}^{\infty}$ (which is constant on the initial BGP) to $\{s_{rt}'\}_{t=0}^{\infty}$ that we will have to solve for numerically. In our analytic results, we take this path as given.

B Calibration of baseline Klette-Kortum model

In our analytical results, we saw that the dynamics of aggregate productivity in the model given a path for the innovation intensity of the economy are shaped by very few parameters. In contrast, to solve the model numerically we must choose all the model's parameters. We now discuss in detail our procedure for doing so based on our baseline Klette-Kortum model (example 3).

Policies: τ_g, τ_e. Given our analytical results, we assume that subsidies to innovation on the BGP in the data are uniform, so that $\tau_e = \tau_g$. As a baseline calibration, we set all policies to zero, but we do indicate in the calibration formulas below where these policies enter.

Interest rate minus growth rate: In our model, on a BGP, the gap between the consumption interest rate and the growth rate of consumption is determined by the consumers' adjusted discount factor $\tilde{\beta} = \beta \exp\left((1-\xi)\,\bar{g}_y\right)$. We set this gap to 0.01, which implies $\tilde{\beta} = 1/1.01$. Given this choice of discount factor, on the BGP, the model interest rate in terms of the research good is given by $r = 0.01$. We set the growth rate of consumption to $g_Y = .02$ and assume an intertemporal elasticity of substitution equal to 1 (i.e. $\xi = 1$), which implies a consumption interest rate of $\bar{r} = 0.03$. This calibration of r is consistent with the difference between the interest rate and the consumption growth rate in McGrattan and Prescott (2012).

Final Consumption Good Production Function: The production function for the final consumption good is parameterized by ρ, which controls the elasticity of the residual demand curve faced by intermediate goods producers. In our baseline calibration we set $\rho = 4$. This elasticity establishes an upper bound on the markup μ that can be chosen.

Equilibrium Markup: Defining *NIPA profits to intangible capital* as $GDP_t - R_{kt}K_t - W_tL_t$, the share of these profits in GDP, denoted by π_t, is given by

$$\pi_t = \left(1 - \frac{1}{\mu}\right) - s_{rt}. \tag{55}$$

From equation (55), the choice of the markup μ is disciplined by data on the innovation intensity of the economy, s_r, and the share of NIPA profits paid to intangible capital relative to GDP. In our baseline calibration, we target a share of NIPA profits paid to intangible capital in GDP, π, of 1% from McGrattan and Prescott (2005) and an innovation intensity of the economy, s_r, of 11% similar to the levels estimated by Corrado et al. (2009) for the United States over the last few years. This implies a markup of 13.6%, $\mu = 1.136$.

Aggregate Production of the Final Consumption Good: Aggregate production of the final consumption good in equation (4) is parameterized by α. We set $\alpha = 0.37$ to match the observation that the share of rental payments to physical capital on the BGP, given by $\frac{\alpha}{\mu}$, is equal to 0.33. With this choice of α, we also have that the share of labor compensation (including production and research) in GDP is given by $\frac{1-\alpha}{\mu} + s_r = 0.66$. The rest of GDP corresponds to profits paid to intangible capital, $\pi = 1\%$.

The Allocation of Labor: The equilibrium allocation of labor between production and research is pinned down in equation (15) by the choices of parameters above and our calibrated innovation intensity of the economy, s_r. In our baseline calibration, with $s_r = .11$, we have $L_p = .833$.

BGP Growth Rates for Scientific Knowledge and Aggregate Productivity: Given our calibration of per capita GDP growth of 2% and our physical capital share of α, we calibrate the growth rate of aggregate productivity in the BGP, $\bar{g}_z$, to 1.25%. For a given choice of γ, the growth rate of

scientific knowledge consistent with these productivity growth rates is given by $g_{Ar} = (1-\gamma)\,\bar{g}_z$. We do not make assumptions about this growth rate directly since we do not observe it. Instead, we alter this parameter as we vary γ.

Innovation Step Size and BGP Innovation Rate: Our choices of innovation step size Δ_z and the BGP innovation rate δ must be consistent with the BGP growth rate of aggregate productivity for intermediate firms, g_z given in equation (12). We must also have that the innovation step size exceeds the markup ($\Delta_z \geq \mu$) to ensure that an incumbent firm has a technological advantage over its latent competitor consistent with its assumed markup. In our baseline calibration, we set $\Delta_z = \Delta_l = \mu$. Given our choice of elasticity ρ and the implied value of $\bar{g}_z$, from equation (12), we get then get $\delta = 0.08$ in our baseline calibration.

Factor $\bar{Y}_r/(\bar{Y}_r - Y_r^0)$: Under the assumptions on policies made above, the ratio $\frac{\bar{Y}_r}{\bar{Y}_r - Y_r^0} = \frac{y_r(1)+\bar{y}_r(0)N(0)}{\bar{y}_r(0)(d(y_r(1))+N(0))}$ is already determined by parameters that we have previously calibrated and by the choice of targets for s_r and π that have been already discussed. To see this, combining the free-entry condition (53) and the following expression for π/s_r,

$$\frac{\pi}{s_r} = \frac{\frac{\Pi Z^{\rho-1}}{\bar{y}_r(0)N(0)P_r} - \left(\frac{y_r(1)}{\bar{y}_r(0)N(0)} + 1\right)}{\left(\frac{y_r(1)}{\bar{y}_r(0)N(0)} + 1\right)}$$

we obtain

$$\frac{y_r(1)}{\bar{y}_r\,(0)\,N(0)} = \frac{\frac{1-\tau_e}{1-\tau_g}\left(1 + \frac{r}{\delta}\left(\frac{d(y_r(1))}{N(0)} + 1\right)\right) - \frac{\xi_g}{1-\tau_g}\left(1 + \frac{\pi}{s_r}\right)}{\frac{\xi_g}{1-\tau_g}\left(1 + \frac{\pi}{s_r}\right) - 1}.$$

Setting $\tau_g = \tau_e = 0$,

$$\frac{\bar{Y}_r}{\bar{Y}_r - Y_r^0} = \frac{\frac{y_r(1)}{\bar{y}_r(0)N(0)} + 1}{\frac{d(y_r(1))}{N(0)} + 1} = \frac{\frac{r}{\delta}}{\xi_g\left(1 + \frac{\pi}{s_r}\right) - 1}.$$

In our baseline calibration, we obtain $\frac{\bar{Y}_r}{\bar{Y}_r - Y_r^0} = 0.966$ so this fraction is close to its upper bound of one, which means that the average cost of innovation by incumbents is close to the marginal cost of innovation by incumbents.

Incumbents' innovation cost function:

Our analytic results showed that the shape of $d\,(.)$ does not matter for the first order aggregate effects of changes in innovation policies (starting from conditional efficiency) except through $\frac{\bar{Y}_r}{\bar{Y}_r - Y_r^0}$, which has already been determined in our baseline calibration as discussed above. For the numerical evaluation of our results (either non-uniform policy changes or relaxing the assumption of conditional efficiency in the initial BGP), we assume that $y_r = \phi_0 + \phi_1 d\,(y_r)^{\phi_2}$, where $y_r \geq \phi_0$. We choose ϕ_0 to target the share of profits in GDP, π, discussed above. We choose the parameter ϕ_1 so that the level of $d\,(y_r\,(1))$ in the initial BGP (from equation (49)) implies a share of employment of entering firms of roughly 3%. In the model, the share of employment by entering firms is given

by

$$s_{new} = \frac{N(0)}{d(y_r(1)) + N(0)} \frac{\exp((\rho - 1)\bar{g_z}) - (1 - \bar{\delta})}{\exp((\rho - 1)\bar{g_z})}.$$

This choice of parameters implies that incumbents account for roughly 75% of total innovations. Our baseline results do not depend on the choice of s_{new} as long as we have positive entry and positive innovation by incumbents. Finally, we follow Acemoglu et al. (2013) and set $\phi_2 = 2.5$. In any calibration of our model, one must check that incumbents actually want to use their innovation technology (i.e. $U > 0$). This is the case in our calibration.

The Intertemporal Knowledge Spillover Parameter γ: Our procedure for calibrating the model to a given BGP does not discipline the parameter γ. In our policy experiments below, we consider three alternative values of γ corresponding to a *low spillover* case ($\gamma = -2$), a *medium spillover case* ($\gamma = 0$) and a *high spillover case* (γ *approaching one*).

Other parameters: The parameters $\bar{y}_r(0)$, σ, and A_r at time 0 can all be normalized to 1 without affecting our results.

C Full numerical solution of baseline Klette-Kortum model (example 3)

In the body of the paper we have presented an analytical partial characterization of the transition dynamics of our model, up to a first-order approximation, following a permanent change in innovation policies. In that approximation, we took as given the transition path for the innovation intensity of the economy and the ratio of physical capital to output. In this section we use the fully calibrated baseline Klette-Kortum model to solve numerically for the full transition dynamics of the economy following a permanent change in innovation policies. We use this solution to evaluate the usefulness of our analytical approximation.

We report results from this experiment in Figures 2 and 3 for versions of our baseline calibrated model with a high, medium, and low intertemporal knowledge spillover as in section 5.

Long run dynamics We first consider the transition dynamics of our model economy over the first 100 years. In Panel A of Figure 2 we show the response of fiscal expenditures as a percent of GDP over the first 100 years of the transition to the new BGP. We see that for each value of intertemporal spillovers that we consider, the change in these fiscal expenditures are very close to their long run value of 3% throughout the transition, independent of the value of intertemporal knowledge spillovers. Thus we see little intertemporal variation in the fiscal impact of the innovation policies that we consider over the long term.

The long run change in the innovation intensity of the economy, s_r, that results from this innovation subsidy is equal to .24 in log terms and is independent of the value of knowledge spillovers. In panel B of Figure 2, we show the dynamics of $\log s_{rt}$ over the first 100 years of the transition.

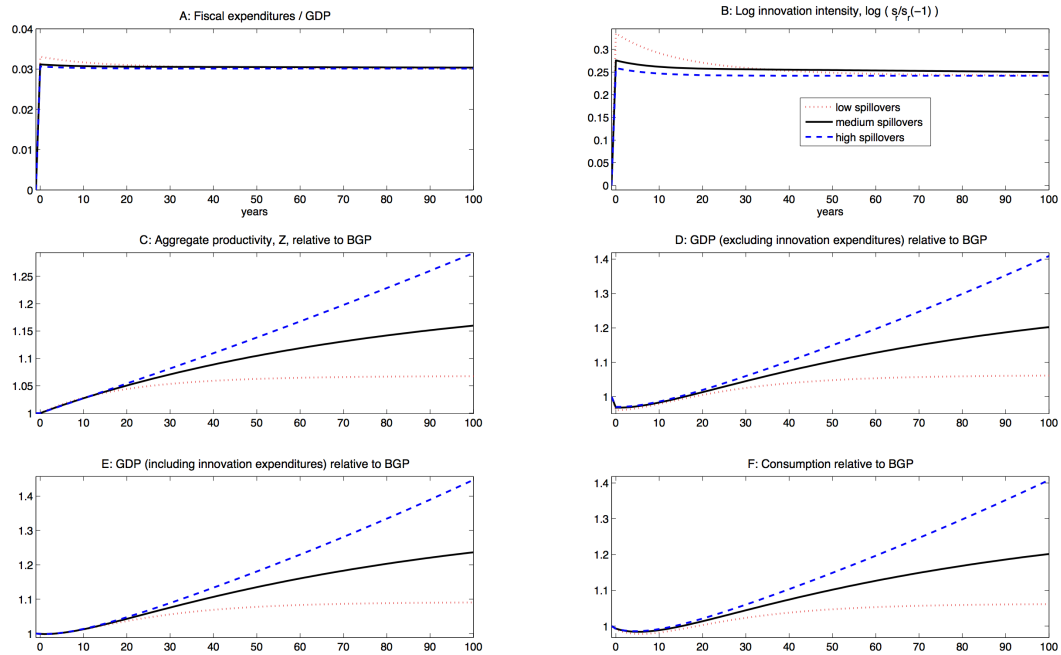

Figure 2: 100-year Transition Dynamics to a Permanent Increase in Innovation Intensity, Full numerical solution, baseline Klette-Kortum model (example 3)

Here we see that there are mild intertemporal substitution effects in innovation expenditures in that this innovation intensity rises by a bit more than .24 in log terms in the early phase of the transition, particularly for the case with low intertemporal knowledge spillovers. The intertemporal substitution of innovation expenditures shown in panel B is more pronounced with low γ because in this case the price of the research good is expected to rise quickly during the transition.

In panel C of Figure 2, we show the response of the level of aggregate productivity during the first 100 years of the transition for our three values of γ. Likewise, in panel D of Figure 2, we show the response of measured GDP (the resources available for consumption and physical investment) during the first 100 years of the transition. The results shown in these panel are quite similar to those for aggregate productivity and measured GDP over the first 100 years shown in Panels E and F of Figure 1. As we will see below, the impact of intertemporal substitution in the path for the innovation intensity of the economy on the transition path for aggregate productivity and of changes in the physical capital to output ratio on measured GDP appear primarily in the initial phase of the transition. In panel E of Figure 2, we show the transition path for measured GDP inclusive of innovation expenditures over the first 100 years. We see that this alternative measure of GDP does not fall substantially on impact as does the measure of GDP excluding innovation expenditures shown in panel D.

In panel F of Figure 2, we show the transition for consumption over the first 100 years. We see

that consumption falls initially (independent of γ) and then rises after a decade or more, where the magnitude of the rise is quite sensitive to γ. As a consequence of these differences in the response of consumption in the very long run, we find that the welfare implications of an increase in innovation subsidies are quite sensitive to the degree of intertemporal knowledge spillovers that we assume. We find in the low spillover case that the equivalent variation in consumption is 3.9%, in the medium spillover case it is 13.9%, and in the high spillover case it is 41.1%. Thus, with large intertemporal knowledge spillovers, we find very large welfare gains from the innovation subsidies in this experiment. The differences in the welfare implications of innovation policy changes across the high and low intertemporal knowledge spillover economies arise as a result of the growing differences in the paths of consumption past the medium term horizon of 20 years in combination with our choice of $\beta = .99$ for the consumer's discount factor (so that the real interest rate is 1 percent higher than the growth rate).

Medium Term Dynamics In Figure 3, we consider the evolution of aggregates during the first 20 years of the transition. We zoom in on this first phase of the transition to better evaluate the quality of our approximation.

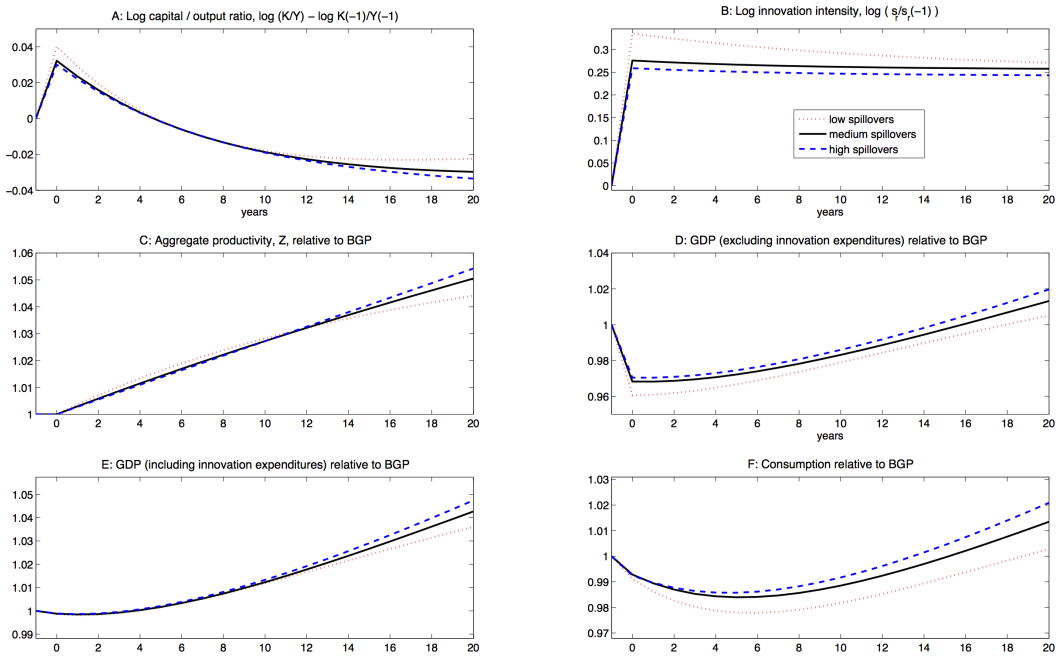

Figure 3: 20-year Transition Dynamics to a Permanent Increase in Innovation Intensity, Full Numerical Solution Klette-Kortum model (example 3)

In panel A of Figure 3, we show the evolution of the log of the physical capital to output ratio over the first 20 years of the transition. This transition path corresponds to the negative of the

transition path for the log of the rental rate on physical capital. In panel B of Figure 3, we show the path of the log of the innovation intensity of the economy. Clearly, there are dynamics of these ratios that we ignored in our analytical approximation shown in panels A and B of Figure 1.

From Proposition 2, we have that the intertemporal substitution in the path for the innovation intensity of the economy shown in Panel B of Figure 3 impacts the model-implied transition for aggregate productivity. By comparing the approximated path for aggregate productivity shown in panel C of Figure 1 to the fully simulated path shown in panel C of Figure 3, we can see that the approximation is fairly accurate despite the dynamics of the innovation intensity of the economy shown in Panel B. In particular, the log of aggregate productivity relative to the initial BGP path in year 20 in the fully simulated path is .053 (versus 0.049 based on the approximation) with $\gamma \to 1$, .049 (versus .044) with $\gamma = 0$, and .043 (versus .035) with $\gamma = -2$.

From expression (42), we have that the full dynamics of GDP are impacted both by the dynamics of the innovation intensity of the economy and of the ratio of physical capital to output. On impact, these two factors have opposite effects on GDP — the initial increase in the ratio of physical capital to output raise GDP while the initial increase in the innovation intensity of the economy above its new long-run level lowers GDP. By comparing the approximated path for GDP shown in panel D of Figure 1 to the fully simulated path shown in panel D of Figure 3, we again can see that the approximation is fairly accurate despite the fact that we ignored the dynamics in the innovation intensity of the economy and the capital-output ratio.

In panels E and F of Figure 3, we show the path of GDP inclusive of innovation expenditures and of consumption during the first 20 years of the transition. We see here that consumption falls initially for at least 10 years during the transition indicating that the potentially large welfare gains are achieved only in the long run.

Medium term dynamics and productivity shocks Figure 3 shows that a large and persistent increase in innovation subsidies has a relatively small impact on aggregate productivity and GDP over a 20 year horizon, and the response of aggregates does not vary much with the extent of intertemporal knowledge spillovers assumed in the model.

These results suggest that it would be hard to verify whether innovation policies yield large output and welfare gains using medium term data on the response of aggregates to changes in innovation policies. We illustrate this point in Figure 4. In that figure we show results obtained from simulating the response of aggregates in our model to our baseline increase in innovation subsidies in an extended version of our model with Hicks neutral AR1 productivity shocks with a persistence of 0.9 and an annual standard deviation of 2%. We introduce these shocks as a proxy for business cycle shocks around the BGP. We show histograms generated from 3000 simulations of the model for the first 20 years of the transition. The units on the horizontal axis show the log of the ratio of detrended GDP at the end of the 20th year of transition to initial GDP and the vertical axis shows the frequency of the corresponding outcome for GDP. In panel A of the figure, we

show results for GDP excluding innovation expenditures. In panel B, we show results for GDP including innovation expenditures. The red bars show results for the model with low intertemporal knowledge spillovers and the blue bars show the results with high spillovers. Clearly, in each panel, the distribution represented by the blue bars is slightly to the right of that represented by the red bars. The histograms in panel B are shifted to the right relative to those in panel A reflecting the fact that GDP including innovation expenditures has a higher elasticity of changes in the innovation intensity of the economy. But it is also clear in each panel that, using either measure of GDP, that it would be very hard to distinguish the degree of intertemporal knowledge spillovers (and, hence, the long term effects from this innovation subsidy) in aggregate time series data even if we had the benefit of a true policy experiment.

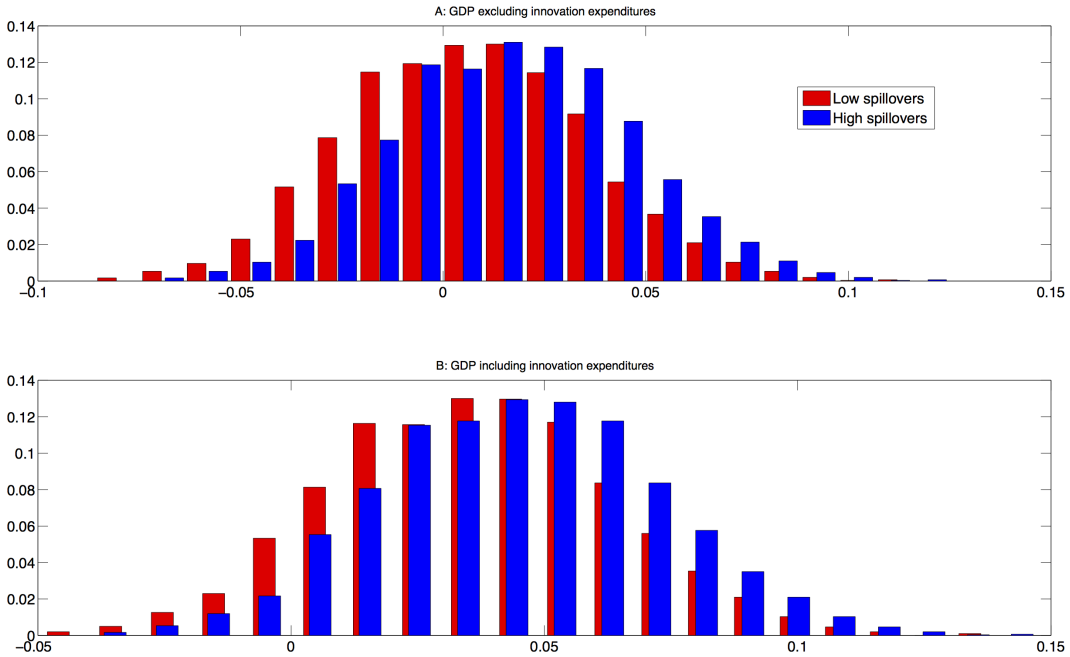

Figure 4: Histogram 20-year Increase in GDP to a Permanent Increase in Innovation Intensity Including Productivity Shocks, Klette-Kortum model (example 3)

Non-uniform changes in innovation policies Up to this point, in our quantitative results, we have considered policy experiments in which the economy starts on a BGP with uniform innovation subsidies and transitions to a new BGP with new uniform innovation subsidies. We use the assumption that the economy initially has uniform innovation subsides to isolate the aggregate implications of changes in innovation policies that work through changing the overall innovation intensity of the economy rather than changing the allocation of research expenditure across firms. We have seen that if the economy starts with uniform innovation subsidies (condi-

tional efficiency), to a first order approximation it does not matter whether the new innovation subsidies are uniform or not.We now solve the full transition dynamics of the model to evaluate whether there are important second order effects that arise when large non-uniform changes in innovation policies are considered. We consider permanent and unanticipated increases in the innovation subsidy to incumbents τ_g only that result in a long run increase in the innovation intensity of the economy from $\bar{s}_r = .11$ to $s_r' = .14$. (In the long run, this subsidy requires fiscal expenditures of 3.3% relative to GDP rather than 3% under our baseline experiment with uniform innovation policies). We show these dynamics for economies with two different curvatures of the incumbents' innovation cost function, $d(.)$, which determines the elasticity of incumbents' innovation, $y_r(1)$, with respect to the incentives to innovate (this elasticity is given by $1/\phi_2$ in our parameterization above). We consider an inelastic case (very high ϕ_2) in which the share of innovation by incumbents, $d(y_r(1))/(d(y_r(1)) + N(0))$, across the old and new BGP is constant at our baseline level of 75%, and an elastic case at our baseline level of $\phi_2 = 2.5$ in which the share of innovative effort by incumbents increases across the old and new BGPs from 75% to 92% (implying a substantial fall in the share of employment in new firms). We set $\gamma = 0$ in all versions of the economy that we consider here. We show results in Figure 5 for the evolution over the first 20 years of the log of the physical capital output ratio, the innovation intensity of the economy, aggregate productivity, GDP exclusive of innovation expenditures, GDP including innovation expenditures, and consumption. In each panel of the figure, we show results from the baseline transition with uniform policies together with the results with inelastic and elastic innovative effort by incumbent firms. Note that the different responses of aggregates between the economies with the non-uniform policy change and inelastic and elastic innovative effort by incumbents arise as a result of second order terms: to a first order approximation, these responses should be the same. The different responses of aggregates between the baseline economy with a uniform policy change (so that the equilibrium $y_r(1)$ does not change) and the economy with a non-uniform policy change and inelastic innovative effort by incumbent firms (so, again, $y_r(1)$ does not change) arise as a result of the different paths for the innovation intensity of the economy due to differences in the intertemporal substitution of innovation expenditure along the transition. We find in both cases that our first order approximation is fairly accurate in that the dynamics of aggregate productivity and measured GDP are not much different than those that we found with uniform changes in innovation policies.

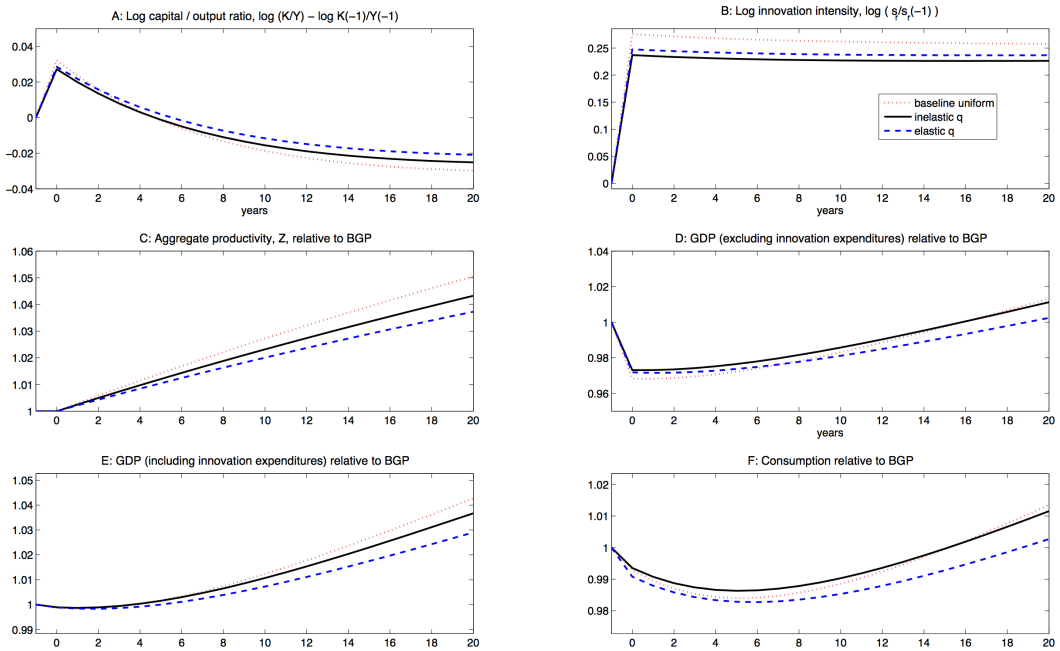

Figure 5: 20-year Transition Dynamics to a Uniform and Non-Uniform Innovation Policy, Medium Knowledge Spillover, Full Numerical Solution Baseline Klette-Kortum model (example 3)

D Optimal allocations and optimal innovation policy in Klette-Kortum model

In Section 4 we used a simple variational argument using the approximated laws of motion for aggregate productivity and GDP derived in Proposition 2 and its corollaries to calculate the optimal innovation intensity of the economy as a function of very few parameters (β, γ, and $\mathcal{E}_g$ in the optimal allocations) for the class of models satisfying our assumptions. The equilibrium of the model depends on additional details of the model such as the market structure and the extent of business stealing (these model details differ between the expanding varieties and quality ladders models). The following proposition characterizes the optimal policy in the baseline Klette-Kortum model (example 3). We allow for an additional policy instrument, a per-unit subsidy on production of the consumption good (denoted by τ_s), that can be set to undo the distortions from the intermediate good's markup. For generality, we also allow for social depreciation of innovation expenditures. The planning problem that we solve in proving this proposition is similar to the one presented in section 4 of Jones (2005).

Proposition 5. *If the social optimum allocation has a BGP with firm-entry and semi-endogenous growth*

($\gamma < 1$), then that optimal BGP corresponds to the equilibrium BGP with a subsidy to the production of the consumption good given by $\tau_s^* = \mu - 1$, and a uniform subsidy to innovation by incumbents and entrants given by

$$\tau_g^* = \tau_e^* = 1 - (\mu - 1)\,\bar{\xi}_g \frac{\frac{r}{\mathcal{E}_g^{KK*}} + 1 - \gamma}{\frac{r}{\bar{\delta}\frac{\bar{Y}_r}{\bar{Y}_r - Y_r^0}} + 1} \tag{56}$$

where $r = \tilde{\beta}^{-1} - 1$, $\bar{\delta}$, $\frac{\bar{Y}_r}{\bar{Y}_r - Y_r^0}$, $\bar{\xi}_g$ and $\mathcal{E}_g^{KK*}$ are all functions of parameters and independent of policies when subsidies to innovation are uniform ($\tau_g = \tau_e$). These policies result in an aggregate allocation of labor in the BGP given by

$$\frac{L_r^*}{L_p^*} = \frac{1}{(1-\alpha)} \frac{\beta \mathcal{E}_g^{KK*}}{\left(1 - \tilde{\beta} + (1-\gamma)\,\tilde{\beta}\mathcal{E}_g^{KK*}\right)}. \tag{57}$$

Note that, as $r \to 0$, the optimal policies and allocation simplify to $\tau_g^* = \tau_e^* = 1 - (\mu - 1)\,(1 - \gamma)$, and $\frac{L_r^*}{L_p^*} = \frac{1}{(1-\alpha)(1-\gamma)}$.

Proof. The result that the optimal innovation by incumbents, $y_{rt}\,(1)$, is common across products follows from the fact that $y_{rt}\,(1)$ enters linearly into the cost of innovation and the returns $d\,(.)$ are strictly concave. Given a common $y_r\,(1)$ across all products, the planner's problem is

$$\max_{\left\{Z_{t+1}, L_{pt}, K_{t+1}, N_t(0), y_{rt}(1)\right\}} \sum_{t=0}^{\infty} \beta^t \left(Z_t K_t^{\alpha} L_{pt}^{1-\alpha} + (1 - d_k)\,K_t - K_{t+1}\right)^{1-\xi}$$

subject the two following constraints

$$-\omega_t\,(1-\xi)\,\tilde{\beta}^t \left\{ \frac{Z_{t+1}}{Z_t} - \left(\sigma(d(y_{rt}\,(1)) + N_t\,(0))\left(\exp(\Delta_z)^{\rho-1} - \zeta\right) + \zeta\right)^{\frac{1}{\rho-1}} \right\}$$

$$-\nu_t\,(1-\xi)\,\tilde{\beta}^t \left\{ \bar{y}_r\,(0)\,N_t\,(0) + y_r\,(1) - A_{rt} Z_t^{\gamma-1}\left(1 - L_{pt}\right) \right\}$$

where ζ is the exponent of the social rate of depreciation. Inspecting this maximization problem, it is clear that the optimal level of $y_{rt}\,(1) = y_r\,(1)$ is the solution to (49) under uniform innovation subsidies, $\tau_g^* = \tau_e^*$ (this is the same condition to obtain conditional efficiency as defined under Assumption 1 above). The FOC w.r.t K_{t+1} is

$$\left(\frac{C_{t+1}}{C_t}\right)^{-\xi} = \beta\left(\alpha \frac{Y_{t+1}}{K_{t+1}} + 1 - d_k\right). \tag{58}$$

Given that the private equilibrium return to capital is $R_{kt} = \frac{\alpha}{\mu}\frac{Y_t}{K_t}$, it follows that $\tau_s^* = \mu - 1$. The F.O.C. w.r.t L_{pt} is

$$\frac{(1-\alpha)}{C_t^{\xi}} \frac{Y_t}{L_{pt}} = \nu_t \frac{Y_{rt}}{1 - L_{pt}}. \tag{59}$$

The F.O.C. w.r.t Z_{t+1} is

$$\omega_t \frac{1}{Z_t} - \omega_{t+1}\tilde{\beta}\frac{Z_{t+2}}{Z_{t+1}^2} = \beta \frac{1}{C_{t+1}^{\xi}} \frac{Y_{t+1}}{Z_{t+1}} + \tilde{\beta}\left(\gamma - 1\right)\nu_{t+1}\frac{Y_{rt+1}}{Z_{t+1}} \tag{60}$$

The F.O.C. w.r.t $N_t\left(0\right)$ is

$$\frac{1}{\rho - 1}\omega_t \left(\frac{Z_{t+1}}{Z_t}\right) \frac{\sigma\left(\exp(\Delta_z)^{\rho-1} - \zeta\right)}{\exp\left(\left(\rho - 1\right) g_{zt}\right)} = \nu_t \bar{y}_r\left(0\right) \tag{61}$$

In the BGP, equation(60) can be written as

$$\exp\left(g_z\right)\omega_t - \omega_{t+1}\tilde{\beta}\exp\left(g_z\right) = \beta\frac{1}{C_{t+1}^{\xi-1}}\frac{Y}{C} + \tilde{\beta}\left(\gamma - 1\right)\nu_{t+1}Y_r.$$

Substituting equation (59) together with the fact that the Lagrange multipliers must be constant in a BGP,

$$\exp\left(g_z\right)\omega\left(1 - \tilde{\beta}\right) = \frac{\tilde{\beta}\nu Y_r}{1-\alpha}\left[\frac{1}{1-\alpha}\frac{L_p}{1-L_p} + \left(\gamma - 1\right)\right].$$

Substituting equation (61),

$$\left(1 - \tilde{\beta}\right) = \tilde{\beta}\frac{1}{\rho - 1}\frac{\sigma\left(\exp(\Delta_z)^{\rho-1} - \zeta\right)}{\exp\left(\left(\rho - 1\right) g_z\right)}\frac{Y_r}{\bar{y}_r\left(0\right)}\left[\frac{1}{1-\alpha}\frac{L_p}{1-L_p} + \left(\gamma - 1\right)\right]. \tag{62}$$

Using the law of motion for Z and $Y_r^0 = y_r\left(1\right) - d\left(y_r\left(1\right)\right)\bar{y}_r\left(0\right)$, we have

$$\sigma\left(\exp(\Delta_z)^{\rho-1} - \zeta\right) = \frac{\exp\left(\left(\rho - 1\right) g_z\right) - \zeta}{\left(Y_r - Y_r^0\right)/\bar{y}_r\left(0\right)}.$$

Substituting this expression into equation (62), we obtain

$$\left(1 - \tilde{\beta}\right) = \tilde{\beta}\mathcal{E}_g^{KK}\left[\frac{1}{1-\alpha}\frac{L_p}{1-L_p} + \left(\gamma - 1\right)\right],$$

where $\mathcal{E}_g^{KK} = \frac{1}{\rho-1}\frac{\exp((\rho-1)\bar{g}_z)-1}{\exp((\rho-1)\bar{g}_z)}\frac{\bar{Y}_r}{\bar{Y}_r-\bar{Y}_r^0}$ as presented in Section 4. This implies equation (57). The innovation subsidy is set so that the socially optimal labor allocation in the BGP equals the equilibrium allocation of labor in the BGP given by

$$\frac{L_p}{1-L_p} = \left(\frac{1-\alpha}{\mu-1}\right)\frac{\Pi Z^{\rho-1}/P_r}{Y_r} = \left(1-\tau_g\right)\left(\frac{1-\alpha}{\mu-1}\right)\frac{\frac{r}{\delta} + \frac{\bar{Y}_r}{\bar{Y}_r-\bar{Y}_r^0}}{\zeta_g\frac{\bar{Y}_r}{\bar{Y}_r-\bar{Y}_r^0}}$$

where the first equality follows from (54) and the second is obtained using (53). □

E Model extension: Occupation choice

Suppose that workers draw a productivity x to work in the research sector, where x is drawn from a Pareto with minimum 1 and slope coefficient $\eta > 1$. There are two wages, W_{pt} and W_{rt}. For the marginal agent,

$$\bar{x}_t W_{rt} = W_{pt}$$

Given that the minimum value of x is 1, any interior equilibrium with positive production requires $W_{rt} \leq W_{pt}$. The aggregate supplies of production and research labor are (having normalized the labor force to 1),

$$L_{pt} = F(\bar{x}_t) = 1\text{-}\bar{x}_t^{-\eta}$$

$$L_{rt} = \int_{\bar{x}_t}^{\infty} x f(x)\, dx = \frac{\eta}{\eta - 1} \bar{x}_t^{1-\eta}.$$

The equilibrium allocation of labor is determined by (as in equation (15) in the baseline model)

$$\frac{W_{pt} L_{pt}}{W_{rt} L_{rt}} = \frac{(1-\alpha)}{\mu_t} \frac{1}{s_{rt}}$$

and

$$\frac{L_{pt}}{L_{rt}} = \frac{\eta - 1}{\eta} \frac{1 - \left(\frac{W_{pt}}{W_{rt}}\right)^{-\eta}}{\left(\frac{W_{pt}}{W_{rt}}\right)^{1-\eta}}$$

Note that as η goes to infinity, W_{pt}/W_{rt} must converge to 1 in order for L_{pt}/L_{rt} to be finite. The elasticity of the aggregate labor allocation with respect to the innovation intensity of the economy (given a constant average markup) is

$$\Delta \log \frac{L_{pt}}{L_{rt}} = -\frac{(\eta - 1)\left(1 + \frac{W_r L_r}{W_p L_p}\right)}{(\eta - 1)\left(1 + \frac{W_r L_r}{W_p L_p}\right) + 1} \Delta \log s_{rt}$$

and the elasticity of research labor with respect to the innovation intensity of the economy is

$$\Delta \log L_{rt} = \frac{(\eta - 1)}{(\eta - 1)\left(1 + \frac{W_r L_r}{W_p L_p}\right) + 1} \Delta \log s_{rt}$$

When η converges to 1 (high worker heterogeneity), the elasticity of L_{pt}/L_{rt} and L_{rt} with respect to s_{rt} converges to 0. When η converges to infinity (no worker heterogeneity), the elasticity of L_{pt}/L_{rt} and L_{rt} with respect to s_{rt} converges to -1 and L_p, respectively, as in our baseline model.

In Proposition 2, equation (36) becomes

$$\log Y'_{rt} - \log \bar{Y}_r = \frac{(\eta - 1)}{(\eta - 1)\left(1 + \frac{W_r L_r}{W_p L_p}\right) + 1} \left(\log s'_{rt} - \log \bar{s}_r\right) - (1 - \gamma)\left(\log Z'_t - \log \bar{Z}_t\right)$$

where $\frac{(\eta-1)}{(\eta-1)\left(1+\frac{W_r L_r}{W_p L_p}\right)+1}$ is bounded between 0 and L_p. In Proposition 2, coefficient Γ_0 is now given by

$$\Gamma_0 = \frac{(\eta - 1)}{(\eta - 1)\left(1 + \frac{W_r L_r}{W_p L_p}\right) + 1} \mathcal{E}_{g0},$$

which is increasing in η (the smaller is the extent of worker heterogeneity, the higher is Γ_0).

F Model Extension: Goods and Labor used as inputs in research

We consider an extension in which research production uses both labor and consumption good, as in the lab-equipment model of Rivera-Batiz and Romer (1991), and discuss the central changes to our analytic results. Specifically, the production of the research good is given by

$$Y_{rt} = A_{rt} Z_t^{\gamma-1} L_{rt}^{\lambda} X_t^{1-\lambda} ,$$

and the resource constraint of the final consumption good is

$$C_t + K_{t+1} - (1 - d_k) K_t + X_t = Y_t.$$

Given this production technology, the BGP growth rate of aggregate productivity g_z is given by $\bar{g}_z = \frac{g_{A_r}}{\theta}$, where $\theta = 1 - \gamma - \frac{1-\lambda}{1-\alpha}$. The condition for semi-endogenous growth is $\theta > 0$. The knife-edge condition for endogenous growth is $\theta = 0$ and $g_{A_r} = 0$, which can hold even if $\gamma < 1$.

Revenues from the production of the research good are divided as follows

$$W_t L_{rt} = \lambda P_{rt} Y_{rt} \text{ , and } X_t = (1 - \lambda) P_{rt} Y_{rt} . \tag{63}$$

The allocation of labor between production and research (the analogous to equation (15) in our baseline model) is related to the innovation intensity of the economy by

$$\frac{L_{pt}}{L_{rt}} = \frac{(1 - \alpha)}{\lambda \mu_t} \frac{1}{s_{rt}} \frac{Y_t}{GDP_t}, \tag{64}$$

where $GDP_t = Y_t - X_t = C_t + K_{t+1} - (1 - d_k) K_t$ when innovation expenditures are excluded in GDP. Factor payments are a constant shares of Y_t.

Our analytical results need to be modified for two reasons. First, the role that the term $1-\gamma$ played in shaping the dynamics of the economy is now played by θ. Second, several of our analytical elasticities need to be modified by the ratio of GDP to Y, which is equal to $\frac{GDP_t}{Y_t} = (1+(1-\lambda)s_{rt})^{-1} \leq 1$.

In Proposition 2, the elasticity of research output Y_r with respect to a change in the innovation intensity of the economy s_r, presented in equation (36), is now given by

$$\log Y'_{rt} - \log \bar{Y}_r = \bar{L}_p \frac{G\bar{D}P}{\bar{Y}} \left(\log s'_{rt} - \log \bar{s}_r\right) - \theta \left(\log Z'_t - \log \bar{Z}_t\right) - \frac{(1-\lambda)\,\alpha}{1-\alpha} \left(\log R'_{kt} - \log \bar{R}_k\right)$$

The third term in the right hand side reflects the change in research output Y_{rt} that result from changes in Y_t relative to K_t when $\lambda < 1$. The coefficients Γ_k in Proposition 2 are now given by

$$\Gamma_0 = \bar{L}_p \frac{G\bar{D}P}{\bar{Y}} \mathcal{E}_{g0} \tag{65}$$

and

$$\Gamma_{k+1} = \left(1 - \theta \mathcal{E}_{g0}\right) \Gamma_k.$$

The result in corollary 1 is now stated as

$$\log Z'_t - \log \bar{Z}_t = \frac{\bar{L}_p \frac{G\bar{D}P}{\bar{Y}}}{\theta} \left(\log s'_r - \log \bar{s}_r\right)$$

Finally, the result in corollary 2 is now adjusted to account for the change in GDP/Y,

$$\log \frac{GDP_t}{Y_t} - \log \frac{G\bar{D}P}{\bar{Y}} = -\left(1 - \frac{GDP}{Y}\right) \left(\log s'_r - \log \bar{s}_r\right).$$