

Optimal Product Variety and Economic Growth: The Trade-off between Internal and External Economies of Scale

Henri L. F. de Groot and Richard Nahuis

Received July 17, 2000; revised version received February 1, 2001

We develop a model to analyze the social optimality of growth and product variety. The model contains two sectors, one assembly sector producing a homogeneous consumption good, and one intermediate goods sector producing differentiated inputs. Both growth which results from R&D performed in the intermediate goods sector and the variety of differentiated intermediate goods are determined endogenously. We disentangle three effects of changes in variety: (i) a productivity effect, (ii) a business stealing effect, and (iii) a growth effect. The market provides too little variety and sub-optimally high growth if the productivity effect of variety is large relative to the market power of intermediate goods producers. If varieties are not very productive, the market provides a too low rate of growth, whereas variety may either be too low as well as too high. Decentralization of both the first-best and a second-best is considered. The model is calibrated to US data to assess the magnitude of required policy interventions and the degree of sub-optimality of growth and product variety. We show that in this framework growth in the market can indeed be sub-optimally high for plausible parameter values.

Keywords: endogenous growth, monopolistic competition, returns to variety, returns to scale, social welfare.

JEL classification: D43, D60, O41.

1 Introduction

Product variety and economic growth are important determinants of economic welfare. In this paper we develop a simple two-sector model of endogenous growth to shed light on the advantages and disadvantages

of using a variety of differentiated inputs. The model builds on recent insights from industrial economics. We explicitly model growth as resulting from intentional research activities. Research builds on tacit knowledge that is specific to firms. In the model *both* the growth rate of product quality and the number of product varieties are endogenous. As a consequence, the model is characterized by a tension between internal and external economies of scale. On the one hand, limited product variety – that is, a small number of large firms – results in strong internal economies of scale and an associated high rate of growth. A large internal scale leaves firms many possibilities to recoup (quasi-fixed) expenditures for growth enhancing activities. On the other hand, large product variety yields (static) external economies of scale by the presence of returns to variety in final-goods production. In this paper, we describe and explore this basic trade-off.

The issue of (optimal) growth and product variety has gathered much attention in the literature.¹ The literature started with the seminal work by Dixit and Stiglitz (1977), and Spence (1976). Dixit and Stiglitz conceive the problem of optimal diversity (in a static context) as one of quantity versus diversity. In the presence of economies of scale in production, producing a small variety saves resources that can be used to expand the production volume. Hence a trade-off arises that leads to the question of social optimality of the market equilibrium. Subsequent studies addressed the optimality question in the presence of growth (e.g., Grossman and Helpman, 1991). In a dynamic context, reduced variety not only saves resources that can be used to extend the produced quantity, but potentially also to increase the rate of growth. The more labor an economy allocates in the R&D sector, the less labor remains for producing consumption goods. The question here is one of growth in quality or variety versus volume of consumption goods. Smulders and van de Klundert (1995) develop an endogenous growth model in which R&D is an in-house activity aimed at improving quality. Besides quality growth, product variety is also determined endogenously. An important characteristic of the model is the constancy of variety in equilibrium.

The model that we develop in this paper differs from the before-mentioned models in several ways. First, and most importantly, we allow for the endogenous determination of both growth and product variety. In

1 In the overview to follow, we have no intention of being exhaustive.

this respect, the model can be interpreted as a combination of the product-variety and quality-ladder model of Grossman and Helpman (1991, chaps. 3 and 4). We model growth as driven by the accumulation of firm-specific knowledge, building on Smulders and van de Klundert (1995), and Peretto (1996). The reasons for pursuing this line are two-fold. First, there is convincing evidence that an important part of knowledge is tacit and hence that knowledge has a firm-specific character, see Dosi (1988). Second, most R&D concerns incremental innovation to adjust existing ideas to firm-specific circumstances. One might also label this the rival activity of knowledge adoption. That is, every firm has to be engaged in costly “re-invention” of any idea from the public knowledge pool before it can be used efficiently. This aspect of knowledge accumulation is a factor of 20 to 30 more important in terms of costs than the creation of non-rival ideas according to Jovanovic (1997), who forcefully makes this point. We incorporate this perspective, and take it to the opposite extreme of the benchmark endogenous growth models, in that R&D is a (purely rival) in-house process for firms. Note that this implies that there is no inter-temporal externality of knowledge. Second, in contrast with Smulders and van de Klundert (1995), we assume that variety has a productivity effect instead of a direct effect on consumers’ welfare (cf. Ethier, 1982; Romer, 1990). We find this approach intuitively more appealing and easier to reconcile with reality (see Barro and Sala-i-Martin, 1995, chap. 6, for a similar argument) and it provides us with the possibility to calibrate the model given the existing empirical evidence on the returns to variety of intermediates (Ciccone and Hall, 1996, and Sect. 5 of this paper). Third, in contrast to the existing literature that studies the tension between growth and variety, we explicitly single out returns to variety (that are external to the firms producing the brands) from the elasticity of substitution between brands. In Dixit and Stiglitz (1977), Grossman and Helpman (1991), and van de Klundert and Smulders (1997), for example, the returns to variety and the elasticity of substitution between brands are mechanically linked. Although the returns to variety and the elasticity of substitution are likely to be (negatively) related, there is no reason to assume a link as tight as suggested in the above-mentioned papers. As we show in this paper, disentangling this relation is crucial for the welfare effects that are derived. In other contexts, characterized by neither endogenous (quality) growth nor endogenous variety, the importance of separating returns to variety from the elasticity of substitution is convincingly shown, see for

example Benassy (1996 and 1998), Broer and Heijdra (2001), de Groot and Nahuis (1998), and Holtz-Eakin and Lovely (1996). These studies put welfare results, derived in the context of models using the traditional Dixit-Stiglitz specification for returns to variety, in perspective. Finally, we calibrate the model to data for the US economy. This provides some evidence for the empirical relevance of the results, especially with regard to the result of potentially sub-optimally *high* growth.

Three distortions are considered in the welfare analysis that we perform. The first distortion is related to the uniqueness of brands giving rise to market power. The resulting non-competitive pricing of differentiated products results in a distortion, as it affects the relative price of inputs in production of the final consumption good. The second distortion arises because entrepreneurs do not take into account the surplus of their entry decision that accrues to the producers of the consumption good, due to increasing returns to variety. The third distortion results from entrants ignoring that entry implies the total market has to be shared by more. The second will be dubbed the *product diversity effect* whereas the third will be called, following Mankiw and Whinston (1986), the *business stealing effect*.² The explicit distinction of these three potential distortions is crucial for understanding the results of the welfare analysis. Crucial is the strength of the product diversity effect relative to the elasticity of substitution. The market results in excessive growth and a sub-optimally low level of variety when the product diversity effect is strong. If the business stealing effect dominates (and hence the product diversity effect is relatively weak), the market growth rate is too low and the variety supplied will be sub-optimally high.³ In the intermediate case, *both* the rate of growth and variety can be too low. Decentralization of the first-best solution is shown to be possible and to require two instruments, namely an output subsidy on the production of intermediates and a lump-sum tax or transfer to producers of intermediates. In addition, we discuss the second-best case in which the government has only one instrument at its disposal, that is, the number of permits it puts on the market for firms in the intermediate sector. Hence, we contribute to the

² The notion of business stealing is analogous to the *profit destruction effect* distinguished by Grossman and Helpman (1991).

³ In an informal discussion on variety and growth, Young (1998) hints at this result in case one deviates from Dixit-Stiglitz preferences.

analyses of optimal growth by analyzing optimal government policy in a model where growth and variety are endogenous (and in the absence of inter-temporal knowledge spillovers). Crucially, returns to variety are distinguished from the elasticity of substitution.

The paper proceeds as follows. In Sect. 2, we discuss and present the basic model. The decentralized steady-state equilibrium is presented in Sect. 3. In Sect. 4, we perform a first-best analysis by determining the unconstrained social optimum, and comparing it with the market outcome. We also design a policy that supports the first-best outcome as a competitive equilibrium. Finally, we provide a second-best analysis in which the government has only one instrument at its disposal. We calibrate the model to US data in Sect. 5. This provides insight in the degree of sub-optimality of growth and product variety and the required policy interventions to reach the first-best outcome. Section 6 concludes.

2 The Model

Our economy comprises two sectors. The assembly sector produces a homogeneous consumption good using intermediates and labor. Firms in this sector operate under perfect competition and take prices of intermediates and wages as given. The intermediates are imperfect substitutes in production of the homogeneous consumption good. Following Ethier (1982), there is an externality in this sector in the form of increasing returns to variety (productivity increases with variety). In the intermediate goods sector, N firms are operating. These firms compete monopolistically. We assume N to be sufficiently large so that competition is monopolistically à la Chamberlin. The number of intermediate goods producers is determined endogenously by a process of entry or exit as long as profits are nonzero.⁴ Intentional R&D performed by firms in this sector results in productivity increases and thus positively affects the quality of intermediates. We assume R&D to be an in-house activity. There is no spill-over of the fruits of R&D whatsoever, i.e., we assume knowledge to be fully tacit.⁵ In this section, we will in turn describe

⁴ The process of entry and exit that is assumed to drive profits to zero can well be extended by explicitly introducing entry costs, see Peretto and Smulders (1998). In our analysis, we abstain from explicitly modeling entry costs; our qualitative results are insensitive to an extension of the model in such a direction.

consumer behavior, the assembly sector, and the intermediate goods sector.

2.1 Consumer Behavior

A representative consumer maximizes his inter-temporal utility subject to a dynamic budget constraint

$$\max \int_0^{\infty} u(C_t) e^{-\theta t} dt \quad \text{s.t.} \quad \dot{A}_t = r_t A_t + w_t L_t - C_t P_{Ct}, \quad (1)$$

where C_t is a consumption index, θ is the subjective discount rate, A_t are assets possessed by consumers,⁶ r_t is the interest rate, $w_t L_t$ is wage income in the economy, and $C_t P_{Ct}$ is expenditure on consumption goods. We assume a constant working population L . Taking $u(C_t) = C_t^{1-\rho}/(1-\rho)$, we arrive at the Ramsey rule

$$\frac{\dot{C}_t}{C_t} = \frac{1}{\rho} \left(r_t - \frac{\dot{P}_{Ct}}{P_{Ct}} - \theta \right), \quad (2)$$

where $1/\rho$ is the inter-temporal elasticity of substitution. So consumers prefer a steeper consumption profile, the larger the gap between the real rate of interest ($r_t - \dot{P}_{Ct}/P_{Ct}$) and the subjective discount rate (θ), and the larger the inter-temporal elasticity of substitution ($1/\rho$).

5 This choice of the knowledge base prevents us from studying interesting issues related to the importance of knowledge spill-overs between firms. It would be fairly straightforward to extend the knowledge base and to introduce knowledge spill-overs. This would, however, significantly complicate the welfare analysis we want to focus on, without adding to the basic insights. It would add two additional market failures, namely (i) an appropriability problem since firms cannot appropriate the full benefits resulting from their own R&D effort due to the spill-overs, and (ii) a problem related to the fact that firms will not take into account the fact that their decision to enter or exit the market will affect the macroeconomic knowledge base on which firms build their research effort. For a welfare analysis related to these issues, we refer to van de Klundert and Smulders (1997).

6 The assets A consist of consumer loans and shares issued by firms in the intermediate goods sector, to finance their investments in research. As, in a consolidated equilibrium, net debt among consumers is zero, the income on the assets, rA , equals the dividends paid by the firms.

2.2 Assembly Sector

The assembly sector produces final consumption goods C under perfect competition. The goods are produced according to (dropping time indices where it leads to no confusion)

$$C = X^\beta L_C^{1-\beta} \quad \text{with } X = N^\sigma \left[\frac{1}{N} \sum_{i=1}^N x_i^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}. \quad (3)$$

In this formulation, L_C represents the number of production workers in the assembly sector. X is a composite of the intermediates that are available. This specification of the composite is borrowed from Broer and Heijdra (2001). The parameter σ captures the returns to variety which is – contrary to Dixit and Stiglitz (1977) – explicitly distinguished from the elasticity of substitution between any pair of different intermediates, ε . This formulation essentially generalizes Dixit and Stiglitz (1977) and emphasizes the fact that the elasticity of substitution and returns to variety are conceptually different, although they are likely to be negatively correlated. In any case, the specific relationship between the returns to variety and the elasticity of substitution as implicitly assumed in Dixit and Stiglitz is arbitrary. We refer to Benassy (1996), Broer and Heijdra (2001), and de Groot and Nahujs (1998) for extensive discussion of this issue (and its consequences for welfare analysis). N represents the number of varieties of the intermediate good (indexed i) available. With all the x_i equal to a common x , as will be the case in equilibrium, we get $X = N^{\sigma-1}(Nx)$. Suppose that there are two bundles of intermediates that are equally large ($N^1 x^1 = N^2 x^2$). It then holds that if $\sigma > 1$ (there are returns to variety), the bundle with the largest variety (the largest N) is most productive. Parameter restrictions $\sigma \geq 1$ and $\varepsilon > 1$ ensure that production exhibits returns to variety (with strict inequality) and that every variety is demanded. The returns to variety equal the special value assumed by Dixit and Stiglitz (1977) if $\sigma = \varepsilon/(\varepsilon - 1)$. It is easily demonstrated that in case $\sigma = \varepsilon/(\varepsilon - 1)$, the expression for the composite good boils down to

$$X = \left[\sum_{i=1}^N x_i^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}, \quad \text{so in symmetry } X = N^{\frac{1}{\varepsilon-1}}(Nx). \quad ^7$$

⁷ This specification is used in many subsequent analyses (for example, Grossman and Helpman, 1991, and van de Klundert and Smulders, 1997).

Producers maximize their profits in two steps (see Appendix A). In the first step, they decide on expenditures on labor and intermediates, taking prices for intermediates, output prices and wages as given.

This yields the standard Cobb-Douglas result of constant cost shares: $XP_X/L_Cw = \beta/(1 - \beta)$. In the second stage, the firm decides on the optimal amount of input of each variety:

$$\max_{x_i} N^\sigma \left[\frac{1}{N} \sum_{i=1}^N x_i^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}} \quad \text{s.t.} \quad \sum_{i=1}^N x_i p_{xi} = XP_X, \quad (4)$$

which results in the downward sloping demand curve for an intermediate of brand i

$$x_i = N^{(\varepsilon-1)(\sigma-1)-1} \left[\frac{p_{xi}}{P_X} \right]^{-\varepsilon} X,$$

where

$$P_X \equiv \frac{\sum_{i=1}^N x_i p_{xi}}{X} = N^{1-\sigma} \left[\frac{1}{N} \sum_{i=1}^N p_{xi}^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}. \quad (5)$$

Hence, under symmetry the true price index ($P_X = N^{1-\sigma} p_x$) is decreasing in N if $\sigma > 1$.

2.3 The Intermediate-goods Sector

The intermediate goods sector consists of N firms, each producing a brand of a differentiated input used in the assembly sector. Firm i produces a quality adjusted amount x_i of the intermediate of type/brand i according to

$$x_i = h_i L_{x_i}, \quad (6)$$

where L_{x_i} represents production labor in the intermediate goods sector, and h_i is the labor productivity. R&D is aimed at quality innovation. The stock of knowledge, which is assumed to be completely firm-specific, directly determines the quality level. Knowledge accumulates according to

$$\dot{h}_i = \xi h_i L_{r_i}, \quad (7)$$

where L_{r_i} represents research labor and ζ is the research productivity parameter.⁸ A dot over a variable denotes time differentiation. Finally, firms need a fixed amount of labor, equal to L_f , before being able to start production. Maximization by firms of their present discounted value subject to the demand for intermediates and the costs and benefits of engaging in R&D gives rise to mark-up pricing, an allocation rule for R&D labor, and the standard no-arbitrage condition (see Appendix A for details).

The model is closed by imposing full employment labor market equilibrium

$$L = L_C + N(L_x + L_r + L_f) , \quad (8)$$

and imposing instantaneous profits in the intermediate goods sector (π) to equal zero

$$\pi = xp_x - (L_x + L_r + L_f)w = 0 . \quad (9)$$

Characterization of the steady-state solution of the model will be the topic of the next section.

3 Solution of the Model

In this section, we characterize the steady-state solution of the model and discuss its main comparative static characteristics (Appendix A provides a formal definition of the steady state and all computational details). A steady state or balanced growth path is defined as a situation with a constant growth rate of labor productivity in the intermediate goods sector and a constant allocation of labor. The model is characterized by the absence of transitional dynamics. We assume that excess profits (or losses) are competed away by free entry and/or exit of firms (ignoring integer constraints). The number of intermediate goods producers is thus determined endogenously. We define the growth rate of labor productivity in the intermediate goods sector as

⁸ For analytical tractability and without affecting the main results, we assume constant returns to R&D labor in the “production” of new knowledge. In an accompanying note (de Groot and Nahuis, 2000), we elaborate on the more general case with diminishing returns to R&D labor.

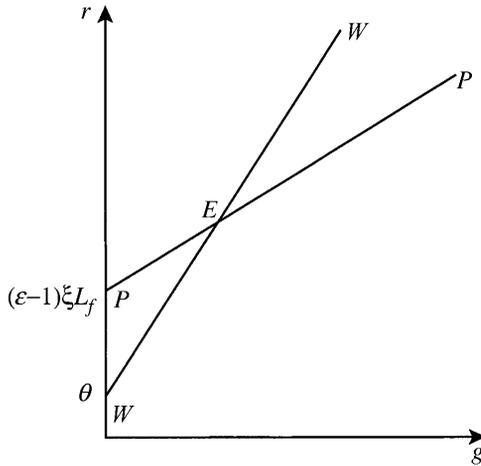


Fig. 1. Equilibrium solution of the model

g ($\equiv \dot{h}/h$). In the remainder of the paper, we take the wage rate as numéraire ($w = 1$).

Solving the model yields a required and realized rate of return on savings and investment, respectively,

$$r = \theta + \beta(\rho - 1)g \quad \text{and} \quad r = (\varepsilon - 1)(g + \xi L_f), \quad (10)$$

where r is the interest rate. The required real rate of return on savings ($r - \dot{P}_C/P_C = r + \beta g$) thus increases with the rate of growth of consumption (βg) due to the wish of consumers to smooth their consumption over time. The realized rate of return depends positively on the rate of growth, the productivity of research and the fixed cost requirement. These two relations are depicted in Fig. 1. The slope of the required rate of return (the WW-locus) equals $\beta(\rho - 1)$ whereas the slope of the realized rate of return (the PP-locus) equals $(\varepsilon - 1)$. Confronting the realized and required rate of return results in a savings-investment equilibrium and yields the equilibrium growth and interest rate

$$g = \frac{(\varepsilon - 1)\xi L_f - \theta}{\beta(\rho - 1) - (\varepsilon - 1)} \quad \text{and} \quad r = (\varepsilon - 1) \frac{\beta(\rho - 1)\xi L_f - \theta}{\beta(\rho - 1) - (\varepsilon - 1)}. \quad (11)$$

In the simple model that we have developed, it makes no sense to think of this equilibrium as being stable or unstable in the usual sense in which equilibrium dynamics take the economy from any initial value of the state variable to the steady state (compare Evans et al., 1998). Since there is only one state variable (h) on which the dynamic equations of the model do not depend, the economy will immediately jump to the steady-state equilibrium. Following Evans et al. (1998), we can investigate the stability properties of the model by going away from the steady state and consider what the dynamics of the economy would be if agents were to learn about the location of the equilibrium. Consider a situation in Fig. 1 in which the growth rate is below the equilibrium growth rate. In such a situation, the realized rate of return on investment (indicated on the PP-locus) exceeds the required rate of return demanded by consumers (indicated on the WW-line). Observing such an out-of-equilibrium situation, firms realize that they can invest in projects with a higher rate of return than required by consumers, whereas consumers realize that they can save more at a higher realized rate of return than required. This results in an adjustment process of expectations that will end when the equilibrium is reached. The opposite logic applies when the initial growth rate exceeds the equilibrium growth rate. This logic applies only when the WW-locus intersects the PP-locus from below. If the WW-locus intersects the PP-locus from above, the before-mentioned logic and adjustment of expectations would take the economy away from the equilibrium and corner solutions would result (see Smulders, 1994). In this “loose” sense we can talk about stability of the equilibrium in the model. Therefore, for stability (that is, $\beta(\rho - 1) > (\varepsilon - 1)$) with a positive growth rate (that is, $(\varepsilon - 1)\xi L_f > \theta$) to apply in the model, we impose the parameter restriction $\beta(\rho - 1) > (\varepsilon - 1) > \theta/(\xi L_f)$.⁹ For a more extensive discussion of stability in this type of model, we refer to Evans et al. (1998). We return to the empirically relevant ranges for these parameter values in Sect. 5.

Finally, we derive the equilibrium number of intermediate goods producers (the variety of intermediates) and the allocation of labor as

$$N = \frac{[\beta(\rho - 1) - (\varepsilon - 1)]\xi\beta L}{\varepsilon[\beta(\rho - 1)\xi L_f - \theta]} \quad \text{and} \quad L_C = (1 - \beta)L. \quad (12)$$

⁹ De Groot and Nahujs (2000) discuss the equilibrium characteristics in a more general setting with a nonlinear research technology. In such a situation, multiple equilibria can arise of which one is stable and the other is unstable (in the sense of stability as discussed in the main text).

Characteristic for the type of model developed in this paper is the constancy of the number of varieties in equilibrium (see Smulders and van de Klundert, 1995). In contrast to, for example, Grossman and Helpman (1991, chap. 3), growth is not driven by increasing variety but instead by increasing quality of the varieties supplied. Obviously, an important element in explaining the non-increasing number of varieties in equilibrium is the fixed cost that is required to produce a new variety (the number of varieties can – by definition – not exceed L/L_f). Furthermore, the model is characterized by the absence of scale effects (that is, the scale of the economy (L) leaves the growth rate unaffected, but leads to an equi-proportionate change in the number of firms). A final important notion is that the parameter capturing the “returns to diversity” does not show up in the solution for the equilibrium variety. The market does not take into account the externality resulting from diversity. The returns to variety parameter is only present in the sense that it affects productivity levels. Hence, it governs the degree of external returns to scale.¹⁰

For the growth rate, we can conclude that it positively depends on the fixed cost requirement L_f and the elasticity of substitution ε . An increase in the fixed-cost requirement lowers the equilibrium number of firms and increases the internal scale of each individual firm. This increases the incentive to engage in R&D and consequently the growth rate. A similar argument applies for the elasticity of substitution. A large elasticity of substitution decreases the room for firms to make positive profits and thus increases individual firm size. An increase in θ and ρ reduces the consumers’ incentive to save and thus increases the costs of acquiring financial means to finance investment in knowledge capital. Firms will respond by investing less in knowledge capital, reducing the rate of growth (and the equilibrium interest rate). The reduction in the investment burden required to keep up with competitors increases the firm’s profits, which will lead to entry.

10 Labor productivity in the assembly sector (C/L_C) grows at rate βg . The labor productivity level equals $\frac{C}{L_C} = (N^{\sigma-1} h \frac{\beta}{1-\beta} \frac{(\varepsilon-1)}{\varepsilon})^\beta$, which positively depends on N (and thus on L) as $\sigma > 1$. This result basically reflects Adam Smith’s notion of division of labor.

4 Optimal Product Variety

In this section, we derive the first-best social optimum (further FBSO) by solving the social planners problem. The social planner maximizes inter-temporal utility of the representative agent solely subject to the technical constraints (i.e., technology and resource availability). Characterization of the social optimum is the topic of Sect. 4.1. In Sect. 4.2, we compare the market equilibrium with the social optimum. The results derived under the Dixit-Stiglitz assumption that $\sigma = \varepsilon/(\varepsilon - 1)$ turn out not to generalize when this assumption is dropped. This emphasizes the importance of disentangling the returns to diversity and the degree of imperfect substitutability resulting in market power of monopolistic competitors. In Sect. 4.3, we design a policy that supports the first-best outcome as a competitive equilibrium. Section 4.4 concludes with a discussion of a second-best world in which the government has only one instrument at its disposal.

4.1 The First-best Social Optimum

In this section, we look at the first-best optimum in which a social planner is assumed to maximize the utility of the representative agent subject to the accumulation function of firm-specific knowledge, the production technology,¹¹ and the labor market constraint. Hence,

$$\max U = \int_0^{\infty} \frac{C^{1-\rho}}{1-\rho} e^{-\theta t} dt, \quad (13)$$

subject to Eqs. (3), (6), (7), and (8). Standard dynamic optimization yields the optimal growth rate (see Appendix B for details about the derivation)

$$g^{\text{FB}} = \xi L_r^{\text{FB}} = \frac{\xi \beta (\sigma - 1)}{\beta (\sigma - 1) + 1} \frac{L}{N^{\text{FB}}} - \xi L_f. \quad (14)$$

¹¹ The model allows for a $\sigma \geq 1$ restriction, but with equality the social planners' problem is not well defined. So here the parameter restriction needs to be narrowed to $\sigma > 1$. It is intuitively clear that it is socially optimal to have a single variety in case $\sigma = 1$. As knowledge is completely firm specific and there is no return to variety as such, it is optimal to minimize on the total fixed cost in the economy. To allow for $\sigma = 1$, the restriction $N \geq 1$ should be added to the social planners optimization program.

The optimal growth rate thus depends negatively on the number of intermediate goods available in the economy. The optimal number of intermediate goods producers can be derived as¹²

$$N^{\text{FB}} = \frac{\beta(\sigma - 1)(\rho - 1) - 1}{\beta(\rho - 1)\xi L_f - \theta} \frac{\beta\xi L}{\beta(\sigma - 1) + 1} ,$$

so
$$g^{\text{FB}} = \frac{\xi L_f - \theta(\sigma - 1)}{\beta(\sigma - 1)(\rho - 1) - 1} . \quad (15)$$

We can derive

$$\frac{\partial N^{\text{FB}}}{\partial \sigma} > 0; \quad \frac{\partial N^{\text{FB}}}{\partial \theta} > 0; \quad \frac{\partial N^{\text{FB}}}{\partial L_f} < 0 . \quad (16)$$

An increase in the returns to variety positively affects the optimal number of differentiated inputs, which is an intuitively clear and important result. An increase in the subjective discount rate increases the optimal number of varieties of the intermediate good. The intuition behind this result is as follows. A higher discount rate increases the value attached to current production. Higher production can be achieved by increasing the number of varieties used in the assembly sector, as the positive productivity effect of an additional variety outweighs the negative growth effect that results from increased fixed cost in the economy. Increases in the fixed cost decrease the optimal product variety. The increased fixed cost tends to lower the optimal rate of growth (if N^{FB} is kept constant). This negative effect can partly be offset by decreasing the fixed-cost requirement by scaling down the number of product varieties.

The above results make clear that the planner faces a trade-off between on the one hand large variety with positive productivity effects in the assembly sector via external returns to scale, and on the other hand high growth rates in the intermediate goods sector via internal returns to scale. The larger the

¹² We impose, almost analogous to the decentralized equilibrium, the parameter restriction $\beta(\rho - 1) > 1/(\sigma - 1) > \theta/\xi L_f$. Appendix B discusses the restriction in greater detail.

returns to variety (the larger σ), the more growth the social planner will sacrifice in exchange for large variety.¹³

4.2 Market Equilibrium Versus the Social Optimum

We are now ready to compare the market outcome as discussed in Sect. 3 with the FBSO. The model is characterized by three potentially distorting market failures. First, there is a (static) distortion in the allocation resulting from the market power that intermediate goods producers have, leading them to engage in mark-up pricing. The second distortion results from the fact that entrants in the intermediate goods sector ignore the productivity effect on the producers of consumer goods (the diversity effect). Third, entrants do not take into account that they decrease the effective market size for their competitors (the business stealing effect).

Figure 2 depicts the comparison between the market equilibrium and the social optimum. Panel (A) depicts the growth rate in the market equilibrium (g^M) and in the social optimum (g^{FB}), whereas Panel (B) depicts the number of firms (product diversity) in the market equilibrium (N^M) and in the social optimum (N^{FB}). The figure is constructed using Eqs. (11), (12), and (15).¹⁴

With respect to the rate of growth, we conclude that the market rate of growth is optimal at $\sigma = \varepsilon/(\varepsilon - 1)$. This is the specific value for the preference for diversity implicitly assumed by Dixit and Stiglitz (1977). The number of product varieties supplied by the market (N^M) is sub-optimally low at this point, which basically restates the Dixit-Stiglitz result in a dynamic environment. Market power of the intermediate goods producers results in non-competitive pricing whereas labor can be hired in a perfectly competitive labor market. Therefore, in the assembly sector, the

13 At this point our model differs crucially from Grossman and Helpman (1991, chap. 3). In their model, growth is growth in product variety. Large returns to variety thus result in sub-optimally low rates of growth (see Benassy, 1998, and de Groot and Nahujs, 1998).

14 From Eqs. (11), (12), and (15), we can derive the partial derivatives of the growth rate and number of firms with respect to σ : $\partial g^M/\partial\sigma = \partial N^M/\partial\sigma = 0$, $\partial g^{FB}/\partial\sigma < 0$, $\partial^2 g^{FB}/\partial\sigma^2 > 0$, and $\partial N^{FB}/\partial\sigma > 0$, $\partial^2 N^{FB}/\partial\sigma^2 < 0$. The point of intersection of g^{FB} and g^M is at $\sigma = \varepsilon/(\varepsilon - 1)$. The point of intersection of N^{FB} and N^M (point q in Fig. 2) is at $\sigma = 1 + [\beta(\rho - 1) + 1]/[\beta\{\rho(\varepsilon - \beta) - (1 - \beta)\}]$. By straightforward calculation it follows that $1 + 1/(\beta(\rho - 1)) < q < \varepsilon/(\varepsilon - 1)$.

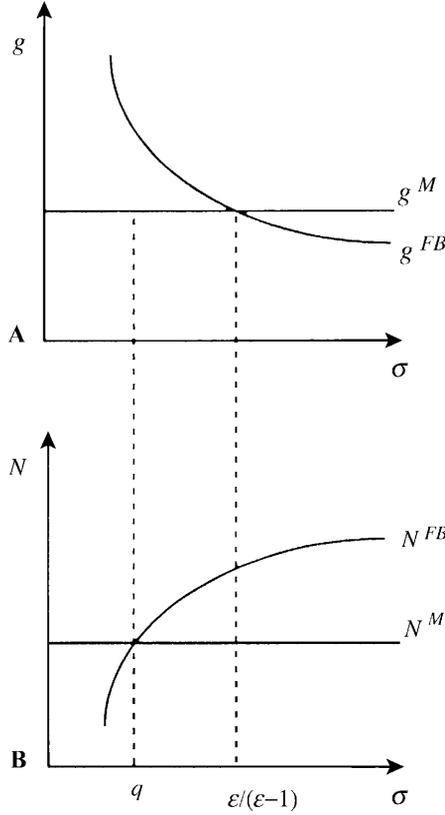


Fig. 2. The market equilibrium versus the social optimum

use of inputs is distorted into the direction of labor (L_C/NL_x is too high from a social point of view). The too low aggregate demand for intermediates results in too little variety of intermediates. The growth rate is at the socially optimal level in the market economy, as in this special case the elasticity of demand *both* reflects excess entry due to business stealing *and* the contribution of variety to overall productivity.¹⁵ Hence, the product

¹⁵ Another way of putting this is to say that firm *size* is optimal, and the decision to invest in R&D is not distorted in any other way. There is no appropriability problem as there are no spill-overs, hence the firm's knowledge base is independent of the number of firms. The private surplus from innovation equals the social surplus.

diversity effect and the business stealing effect happen to be of equal magnitude (*cf.* Grossman and Helpman, 1991, Appendix A.3.3).

To the right of the intersection of g^M and g^{FB} , the returns to variety are strong but *not* (fully) reflected in the market by a high mark-up (as in the special case discussed above). The relatively low mark-up means that profit opportunities are gloomy and hence the market supports only a low number of firms (see panel B of Fig. 2). The number of firms in the market is sub-optimally low. Investing in R&D is a fixed cost that can be spread over more sales if the market is larger. Hence, the existence of a low number of firms makes the return to R&D high and causes the market to produce a sub-optimally high growth rate. To the left of the intersection of the curves N^M and N^{FB} , the argument presented above can be reversed; when the returns to variety are low, the market provides too much variety and a too low rate of growth. In the intermediate case where $q < \sigma < \varepsilon/(\varepsilon - 1)$ the market provides too little variety and too low a rate of growth. The intuition for this result can be explained in three steps. (1) In the parameter range under consideration the social return to variety is relatively low. The entry decision, however, is based on the price elasticity (ε) that is relatively low (profit opportunities are relatively high). This suggests a tendency for excessive entry. (2) A countervailing power arises by distortionary pricing of differentiated goods producers. Cost minimization of the assembly firms leads to too low aggregate demand for intermediates. This downwardly biases the entry decision. On balance entry is sub-optimally low. (3) As we have seen, non-competitive pricing as such does not impede a Pareto optimal growth rate (see panel A at $\sigma = \varepsilon/(\varepsilon - 1)$). For the growth rate we only have to take into account the part of the argumentation under (1).¹⁶ This implies, following the logic phrased above, a sub-optimally low rate of growth.

4.3 Decentralizing the First-best Social Optimum

Having determined the first-best social optimum, we are now ready to analyze how a policy should look like that can support the first-best social

¹⁶ This reasoning can be illustrated as follows: if mark-up pricing (leading to too low aggregate demand for intermediates) would evaporate, the distortions on growth and on variety would be balanced again. This is easily demonstrated by setting β equal to one (i.e., no direct labor is used in the production of consumption goods). It then holds that $q = \varepsilon/(\varepsilon - 1)$, so the area between q and $\varepsilon/(\varepsilon - 1)$ in Fig. 2 vanishes.

optimum as a competitive equilibrium. Having three distortions in our model, three instruments are in principle required to decentralize the FBSO. However, as we show, in this specific model two instruments suffice.¹⁷ We introduce two policy instruments, namely an output subsidy on intermediates (s) and a lump-sum tax levied on firms producing the intermediates (T).¹⁸ We assume that the revenues that the government needs, in order to be able to pay the subsidies, are generated in a non-distortionary lump-sum fashion by taxing consumers. Instantaneous profits of firms producing intermediates now equal $\pi = xp_x(1 + s) - (L_x + L_r + L_f)w - T$. Following the same procedure as in Sects. 2 and 3, we can solve for the equilibrium rate of growth and interest, the equilibrium number of intermediate goods producers, and the equilibrium allocation of labor over the various activities. Putting the equilibrium values equal to those in the first-best social optimum yields the optimal output subsidy and lump-sum tax. This gives rise to

$$s = \frac{1}{\varepsilon - 1} > 0 \quad \text{and} \quad T = \left[\frac{\varepsilon}{\varepsilon - 1} - \sigma \right] \frac{\xi L_f \beta (\rho - 1) - \theta}{\xi [\beta (\sigma - 1) (\rho - 1) - 1]} . \quad (17)$$

Decentralization of the first-best social optimum thus requires an output subsidy that is larger the more market power producers of intermediates have (i.e., the less substitutable the intermediates are). This subsidy corrects for the distorted price of the intermediates due to monopolistic pricing and the associated too low demand for intermediates that results from this price distortion. The lump-sum tax (or subsidy, depending on whether the returns to variety as measured by σ is larger or smaller than the mark-up) is required to correctly size firms or, to put it alternatively, to sustain the first-best number of firms. It thereby corrects for the profit destruction effect and the product diversity effect. Both the size and sign

¹⁷ This result resembles the result in Barro and Sala-i-Martin (1995, chap. 6) where in a model more or less similar to ours (without explicitly distinguishing between returns to variety and the elasticity of substitution) static and dynamic inefficiency can be eliminated with a single instrument. A similar issue arises in Grossman and Helpman (1991, chap. 3) where three distortions can be eliminated using one instrument.

¹⁸ We could also introduce a subsidy on the use or purchasing of intermediates in the assembly sector. It is easily verified that this subsidy should equal $1/\varepsilon$. This subsidy is required to correct for the distorted price of the intermediates due to monopolistic price setting.

of the lump sum tax depend crucially on the returns to variety and the elasticity of substitution between intermediates. If returns to variety are relatively large compared to the mark-up ($\sigma > \varepsilon / (\varepsilon - 1)$), a subsidy should be granted to the producers of intermediates ($T < 0$) in order to keep more firms alive than could otherwise be sustained. This is understandable since the problem of excess entry resulting in profit destruction is limited given the limited market power firms have, while the problem of insufficient entry resulting from the returns to variety is large. Hence, there is insufficient entry and a subsidy is required. When returns to variety are relatively small, a lump-sum tax should be levied. In the special case where the returns to diversity are exactly equal to the mark-up, no lump-sum transfer is needed and the first-best social optimum can be achieved by just granting the producers of intermediates an output subsidy.

4.4 A Second-best Social Optimum

Successful implementation of industrial policies to achieve the first-best results derived in the previous section may be difficult and unrealistic. They require enormous amounts of information available to the government (and the government to behave in a benevolent manner). In this section, we therefore consider the second-best problem of choosing the welfare maximizing number of firms, taking as given the non-competitive behavior after entry of intermediate goods producers. Hence, it is assumed that the social planner cannot control the behavior of a given number of firms, but can choose the number of allowances. In order to derive the socially optimal number of permits to be put on the market, we derive an explicit expression for welfare (given the behavior of firms).

Recall the utility function introduced in Sect. 2.1 ($U_0 = \int C_t^{1-\rho} / (1-\rho) e^{-\theta t} dt$). Using $C_t = X_t^\beta L_{C_t}^{1-\beta}$ and the fact that the allocation of labor and the growth rate of productivity and consumption are constant over time, we arrive at

$$U_0 = \int_0^\infty \frac{\left[(N^\sigma h_0 L_x)^\beta L_C^{1-\beta} \right]^{1-\rho}}{1-\rho} e^{[(1-\rho)\beta g - \theta]t} dt, \quad (18)$$

where h_0 is the initial productivity level (which subsequently grows at the constant rate g). Integrating this expression yields

$$U_0 = \frac{-1}{(\rho - 1)[\beta g(\rho - 1) + \theta] \left[(N^\sigma h_0 L_x)^\beta L_C^{1-\beta} \right]^{\rho-1}}. \quad (19)$$

The optimal amount of allowances to be put on the market is then easily determined by deriving the sign of the derivative of the present discounted utility w.r.t. N which equals

$$\begin{aligned} & \text{sgn.} \left(\frac{dU_0}{dN} \right) \\ &= \text{sgn.} \left(\sigma \beta \frac{1}{N} + \beta \frac{1}{L_x} \frac{\partial L_x}{\partial N} + (1 - \beta) \frac{1}{L_C} \frac{\partial L_C}{\partial N} + \frac{\beta}{\beta g(\rho - 1) + \theta} \frac{\partial g}{\partial N} \right). \end{aligned} \quad (20)$$

Utility is affected by an increase in the number of varieties of the intermediates (i.e., the number of permits issued by the government) through four channels:¹⁹

- (1) There is a direct *positive* productivity effect (the first term); this is due to the diversity effect in the Ethier specification;
- (2) There is a *negative* volume effect (second term) related to business stealing or profit destruction; more varieties reduce the size of firms producing intermediates and thereby reduce the produced volume of intermediates;
- (3) There is a *negative* effect on production of consumption goods (third term); more varieties are resource consuming due to the increase of the total fixed cost in the economy, reducing labor available for production of the consumption good in the assembly sector;
- (4) There is a *negative* growth effect (the fourth term); more varieties reduce the scale of operation for each firm and hence the profitability of engaging in R&D.

The negative effects are essentially caused by two factors, *viz.*, an effect working via the “potential of the economy to have large intermediate goods producers with a high growth potential” (L/N), and an effect

¹⁹ Note that in performing this analysis, we look at an *exogenous* change in N . The number of firms is determined exogenously. We assume in other words that entry is blocked. The solution of the model under the assumption that there is no free entry is given in Appendix C.

working via the supply of labor available for productive purposes ($L - NL_f$) that is affected by increases in fixed costs following an increase in variety.

The optimal number of firms in the second-best social optimum (SBSO) can now be derived as

$$N^{\text{SB}} = \frac{\beta(\sigma - 1)(\rho - 1) - 1}{\beta(\rho - 1)\xi L_f - \theta} \frac{\beta \xi L}{\beta(\sigma - 1) + 1}. \quad (21)$$

Comparing this solution with the solution for the FBSO (Eq. (15)) reveals that the product variety chosen by the planner who has only one instrument at his disposal is exactly equal to the FBSO. This reflects the fact that the marginal utility of an additional variety is independent of the rate of growth. So comparing the SBSO for the number of intermediates with the solution in the decentralized equilibrium gives rise to the same conclusions as described in Sect. 4.2. The growth rate in the second best is lower than in first best. Comparing the growth rate in the second best with the market equilibrium yields the same qualitative conclusions as in Sect. 4.2. In the special case of $\sigma = \varepsilon/(\varepsilon - 1)$, the growth rate implicitly chosen by the planner is lower than the market growth rate (remember that the planner chooses the first-best number of firms, which is higher than the market supports), while growth is sub-optimally high for low returns to variety and vice versa. This reflects the predetermined weight on the lack of variety in the Dixit-Stiglitz specification, for which the planner solves. In the special case where $\beta = 1$ (i.e., only intermediates are used in the assembly sector and hence the potentially distorting market failure due to mark-up pricing becomes ineffective as relative prices in the assembly sector are no longer distorted), the second-best growth rate equals the first best.²⁰ In other words, when the distortion in factor allocation resulting from mark-up pricing by intermediate goods producers is “eliminated”, the social planner who has only one instrument at his disposal (the number of permits) can perfectly replicate the FBSO.

5 A Calibration Exercise

In this section, we again look at the FBSO discussed in the previous section. The aim of this section is to calibrate the model in order to assess the empirical

²⁰ We refer to Appendix C for the solution of the growth rate.

relevance of the result that growth can be too high in a decentralized market equilibrium. Second, we seek to determine the size of the policy interventions required to arrive at the FBSO.

To do so, we use values for the endogenous variables in the model as reported in Table 1. These values represent the US economy in 1990. The growth rate of labor productivity in the manufacturing sector is the average over the period 1960–1995 taken from the International Sectoral Database (OECD, version 1998.1).²¹ We take the share of fixed labor in total intermediate-goods sector employment from the actual data, as the fixed costs are crucial in determining the optimal firm size in our model. For consistency, we also take the other employment shares from the data. These employment shares of different labor-types in the intermediate-goods sector are derived from various issues of the *Year Book of Labour Statistics* from the International Labour Organization.²² By choosing to use employment shares from the data, the mark-up is determined, since the share of production labor in total employment is inversely related to the mark-up (see Eq. (A.13) in Appendix A). The inter-temporal elasticity of substitution is taken from Hall (1988). Total employment is taken as the sum of manufacturing employment and service employment (ISIC 3 and 5–8). The share of the intermediate goods sector is taken from the ISDB and equals the share of manufacturing employment (ISIC 3) in total employment. The parameter capturing returns to diversity is derived from Ciccone and Hall (1996).²³

Starting from these values, we calculate the mark-up, the subjective discount rate, the interest rate, and sectoral employment. The mark-up

²¹ The time average is taken to avoid cyclical influences.

²² They classify manufacturing labor in seven sub-groups, namely (1) professional, technical and related workers, (2) administrative and managerial workers, (3) clerical and related workers, (4) sales workers, (5) service workers, (6) agricultural workers and (7) production workers. Group 1 is interpreted as R&D labor, groups 2–6 as overhead or fixed labor and group 7 as production labor. The shares of these groups in total manufacturing employment are given in Table 1.

²³ Ciccone and Hall empirically explain differences in labor productivity across US states by differences in the diversity of local intermediate services. The parameter μ in their model, which captures the substitutability of intermediate products, reflects our parameter measuring the returns to diversity. Assuming the degree of diminishing returns due to the presence of a fixed factor (i.e., land) which is captured by α in their model to be equal to 0.85, the returns to diversity can be determined as equal to 1.88 (details are available upon request from the authors).

Table 1. Calibration values

Growth of labour productivity	g	0.0273
Fixed-cost share	$L_f/(L_x + L_r + L_f)$	0.29
Production share	$L_x/(L_x + L_r + L_f)$	0.59
R&D share	$L_r/(L_x + L_r + L_f)$	0.12
Inter-temporal elasticity of substitution	$1/\rho$	0.1
Share of intermediate goods sector	β	0.26
Returns to diversity	σ	1.88
Total employment (millions)	L	94

follows from the fact that the zero-profit condition implies that the mark-up equals $(L_x + L_r + L_f)/L_x$. The elasticity of substitution between intermediates is thus equal to 1.70. This value falls inside the range of elasticities found in the literature (for example Hall, 1988, and Roeger, 1995). Using Eq. (11) and the fact that $r = gL_x/L_r$, the subjective discount rate θ is then straightforwardly derived as 0.07; a value close to the “common sense estimate”. The sectoral allocation of labor follows from the fact that $L_C = (1 - \beta)L = 69.7$ million.²⁴

Having calibrated the model, we can now determine the tax and subsidy required to achieve the FBSO. This is done for different values of σ (the lower and upper bound used here are two standard errors away from the estimated value reported in Ciccone and Hall, 1996). The results are reported in Table 2.

A first important result is that for reasonable parameter values, growth in the market is too high from a planner’s perspective. For the range of parameter values reported in the table, the number of firms sustained in the FBSO in all cases exceeds the number of firms sustained in the market equilibrium (see also Fig. 2). In interpreting this result, it is important to note that the model neglected a potentially important externality that has been emphasized in the literature and that tends to result in the market growth rate being too low, namely knowledge spillovers. Therefore, the calibration results only show that in the specific context of this model, it is *possible* that growth in the market is excessively *high* when accounting

²⁴ We can also determine ζL_f – crucial for the stability condition – which is straightforwardly determined from Eq. (15) as 0.066 (note that ζ is a pure scale parameter that is meaningful only in relation to the size of the high-tech firm). It is easy to verify that the inequality restrictions imposed in Sect. 3 hold for this parameter constellation.

Table 2. Policy results for different values of the taste of variety (σ)

	$\sigma = 1.81$	$\sigma = 1.88$	$\sigma = 1.94$	Market
Subsidy rate (s)	1.434	1.434	1.434	0
Lump-sum tax (T)	-5.771	-7.766	-9.24	0
Tax as % of sales (T/xp_x)	-0.047	-0.06	-0.075	0
First-best growth rate (g^{FB})	0.01	0.005	0.0003	0.027
$N^{\text{FB}}/N^{\text{M}}$	1.991	2.297	2.592	

for the productivity externality of diversity. Keeping this in mind, we can turn to the second result from the calibration exercise, which is the assessment of the required policy interventions to reach the FBSO. The table indicates that at the mean value of the parameter capturing the returns to diversity, a lump sum subsidy of 6.3% of sales is sufficient to raise the number of firms to first-best level. Obviously, the required lump-sum transfer declines as the returns to diversity decline (and vice versa). This lump sum subsidy is to be accompanied by a subsidy rate of 143% to correct for the distortion of mark-up pricing. As the returns to variety leave the optimal internal allocation of labor within the firm unaffected, this subsidy rate does not depend on the degree of returns to variety.

6 Conclusion

This paper analyzed the social optimality of growth and product variety. To do so, we developed a model that is characterized by the endogenous determination of *both* growth and product variety. In our model, a trade off exists between internal and external economies of scale. The trade off arises since product variety on the one hand enhances productivity due to the presence of the returns to variety, but on the other hand leaves a smaller internal scale and hence less room for growth-enhancing R&D. By explicitly singling out the elasticity of substitution and the returns to product variety, we have shown that the market does not always provide an equilibrium that reflects a socially desirable trade-off between variety and growth. When returns to diversity are strong relative to the elasticity of substitution between varieties of the high-tech good, external economies of scale are important and the market supports a too high growth rate and yields a lack of variety. When the returns to diversity are very weak and internal economies of scale are important, this conclusion is reversed.

In the intermediate case, both the growth rate and the number of varieties are too low from a social point of view. For reasonable parameter values, excessively high growth was shown to be potentially relevant. These results illustrate the importance of disentangling the returns to variety and the imperfect substitutability of brands when analyzing the welfare properties of models of monopolistic competition with endogenous growth and endogenously determined market structures.

Appendix

A Solution of the Model

In this Appendix, we characterize the steady-state solution of the model. Firms in the assembly sector perform a two-stage maximization procedure. In the first stage, they solve

$$\max_{L_p, X} \Pi_C = CP_C - wL_C - P_X X, \quad (\text{A.1})$$

where Π_C denotes profits and P_X is the price index for the composite good. Optimization yields

$$\frac{\partial \Pi_C}{\partial X} = 0 \Leftrightarrow \beta CP_C = XP_X, \text{ and} \quad (\text{A.2})$$

$$\frac{\partial \Pi_C}{\partial L_C} = 0 \Leftrightarrow (1 - \beta)CP_C = L_C w, \quad (\text{A.3})$$

which is the standard Cobb-Douglas result of constant cost shares.

The objective of firms in the intermediate goods sector can be written as

$$\max \int_0^\infty [x_i p_{x_i} - (L_{x_i} + L_{r_i} + L_f)w] e^{-rt} dt. \quad (\text{A.4})$$

r is the interest rate at which firms could invest their money in the financial market. The current-value Hamiltonian corresponding to this dynamic optimization problem reads as

$$H = x_i p_{x_i} - (L_{x_i} + L_{r_i} + L_f)w + p_{h_i} \xi h_i L_{r_i}, \quad (\text{A.5})$$

where p_{h_i} is the shadow price corresponding to knowledge. The first-order conditions to the inter-temporal optimization problem are (assuming symmetry so that we can drop the brand index i)

$$\frac{\partial H}{\partial L_x} = h p_x \left(1 - \frac{1}{\varepsilon} \right) - w = 0 \Leftrightarrow p_x = \frac{\varepsilon}{\varepsilon - 1} \frac{w}{h}, \quad (\text{A.6})$$

according to which firms engage in mark-up pricing,

$$\frac{\partial H}{\partial L_r} = -w + p_h \zeta h = 0 \Leftrightarrow w = p_h \zeta h, \quad (\text{A.7})$$

showing that firms allocate R&D labor to this sector as long as the marginal benefits of doing so ($p_h \zeta h$) exceed the marginal costs (w), and

$$-L_x p_x \frac{\varepsilon - 1}{\varepsilon} - p_h \zeta L_r = \dot{p}_h - r p_h \Leftrightarrow \frac{\dot{p}_h}{p_h} + L_x \frac{p_x}{p_h} \frac{\varepsilon - 1}{\varepsilon} + \zeta L_r = r, \quad (\text{A.8})$$

which is the no-arbitrage condition. Using the expressions for the price of intermediates (Eq. (A.6)) and the allocation rule of research labor (Eq. (A.7)), we can derive

$$\frac{\dot{p}_x}{p_x} = \frac{\dot{w}}{w} - \frac{\dot{h}}{h} = \frac{\dot{w}}{w} - g = \frac{\dot{p}_h}{p_h}. \quad (\text{A.9})$$

Note that the wage rate is the numéraire ($w = 1$). In the steady state, the allocation of labor is constant, and the number of firms is fixed. This gives rise to

$$g_C = \beta g_x = \beta g \quad \text{and} \quad \frac{\dot{P}_C}{P_C} = \beta \frac{\dot{p}_x}{p_x} = -\beta g. \quad (\text{A.10})$$

So we can write the Ramsey rule as

$$r = \theta + \beta(\rho - 1)g. \quad (\text{A.11})$$

This equation corresponds to Eq. (10) in the main text.

Using Eqs. (A.6), (A.7), and (A.9), the no-arbitrage condition (A.8) yields

$$r = \xi L_x . \quad (\text{A.12})$$

Imposing free entry and exit in the intermediate goods sector results in zero-excess profits (Eq. (9)). Substituting the price for intermediates (Eq. (A.6)) and Eq. (6) into the zero-profit condition (Eq. (9)), it boils down to

$$\frac{L_x + L_r + L_f}{L_x} = \frac{\varepsilon}{\varepsilon - 1} \Leftrightarrow L_x = (\varepsilon - 1)(L_r + L_f) . \quad (\text{A.13})$$

The firm size in relation to the size of the production department is thus equal to the mark-up. The rate of return on investment can now be written as

$$r = (\varepsilon - 1)[g + \xi L_f] , \quad (\text{A.14})$$

which is Eq. (10) in the main text.

The equilibrium number of high-tech firms and the allocation of labor can then be determined using labor market equilibrium (Eq. (8)), and

$$CP_C = \frac{Np_{x_i}x_i}{\beta} = \frac{L_C w}{1 - \beta}, \quad \text{so} \quad \frac{NL_x \frac{\varepsilon}{\varepsilon - 1}}{\beta} = \frac{L_C}{1 - \beta} , \quad (\text{A.15})$$

which shows how the assembly sector optimally chooses between labor and intermediates. This leaves us with

$$N = \frac{[\beta(\rho - 1) - (\varepsilon - 1)]\xi\beta L}{\varepsilon[\beta(\rho - 1)\xi L_f - \theta]} \quad \text{and} \quad L_C = (1 - \beta)L . \quad (\text{A.16})$$

This equation corresponds to Eq. (12) in the main text.

B The First-best Social Optimum

This Appendix solves for the first-best social optimum. We allow for diminishing returns to research labor. The current-value Hamiltonian corresponding to the optimal-control problem given in Sect. 4.1 is²⁵

²⁵ Use that under symmetry $X = N^{\sigma-1}Nx = N^\sigma hL_x$.

$$H = \frac{1}{1-\rho} \left[(L - N(L_x + L_r + L_f))^{1-\beta} (N^\sigma h L_x)^\beta \right]^{1-\rho} + p_h \xi h L_r . \quad (\text{B.1})$$

The first-order conditions corresponding to this problem are

$$\frac{\partial H}{\partial L_x} = 0 \Leftrightarrow (1-\beta)NL_x = \beta L_C , \quad (\text{B.2})$$

$$\frac{\partial H}{\partial L_r} = 0 \Leftrightarrow p_h = (1-\beta) \frac{C^{1-\rho} N}{L_C \xi h} , \quad (\text{B.3})$$

$$\frac{\partial H}{\partial N} = 0 \Leftrightarrow L_C = \frac{(1-\beta)L}{\beta\sigma + 1 - \beta} = L_C^{\text{FB}} , \quad (\text{B.4})$$

$$\dot{p}_h + \frac{\partial H}{\partial h} = \theta p_h \Leftrightarrow \dot{p}_h + \frac{C^{1-\rho} \beta}{h} + p_h \xi L_r = \theta p_h . \quad (\text{B.5})$$

From Eq. (B.4), it is evident that employment in the assembly sector is constant. Using Eq. (B.2), it then follows that NL_x is constant. Substitution of the expressions for NL_x and L_C in the labor market constraint (Eq. (8)) gives an expression for L_r and consequently for the optimal growth rate as given in the main text.

Next, we turn to the determination of the optimal number of firms in the intermediate goods sector in the first-best social optimum. Differentiation of Eq. (B.3) with respect to time yields

$$\frac{\dot{p}_h}{p_h} = (1-\rho) \frac{\dot{C}}{C} + \frac{\dot{N}}{N} - \frac{\dot{L}_C}{L_C} - \frac{\dot{h}}{h} . \quad (\text{B.6})$$

We know by combining Eqs. (B.3) and (B.5) that

$$\frac{\dot{p}_h}{p_h} = \theta - \xi L_r - \frac{\beta \xi L_C}{N(1-\beta)} . \quad (\text{B.7})$$

By differentiation of Eq. (3) with respect to time, we get (using symmetry among the intermediate goods producers)

$$\begin{aligned} \frac{\dot{C}}{C} = \beta \frac{\dot{X}}{X} + (1 - \beta) \frac{\dot{L}_C}{L_C} = \beta \left[(\sigma - 1) \frac{\dot{N}}{N} + \frac{\dot{N}}{N} + \frac{\dot{h}}{h} + \frac{\dot{L}_x}{L_x} \right] \\ + (1 - \beta) \frac{\dot{L}_C}{L_C} . \end{aligned} \quad (\text{B.8})$$

We now use the fact that in the steady state the allocation of labor is constant and that labor-market equilibrium holds to arrive at

$$\frac{\xi(L_r + L_f)}{(\sigma - 1)} = \theta + \beta(\rho - 1)\xi L_r . \quad (\text{B.9})$$

This equation reflects the fact that in the optimal allocation, the costs of investing in R&D (the right-hand side of Eq. (B.9)) should equal the benefits (the left-hand side of Eq. (B.9)).

C The Blocked Entry Model (N fixed)

This Appendix solves the model under the assumption that the number of firms is fixed ($N = \bar{N}$). The system of equations from which the solution can be found consists of Eqs. (A.11), (A.12), (A.15), (7) (combined with the definition for the rate of growth), (8), and $N = \bar{N}$. We can solve this system of six equations with six unknowns (g, r, L_x, L_r, L_C, N).

The solution for the growth rate is

$$g = \frac{\xi \frac{L}{N} - \xi L_f - \frac{\varepsilon - \beta}{\beta(\varepsilon - 1)} \theta}{(\rho - 1) \frac{\varepsilon - \beta}{\varepsilon - 1} + 1} . \quad (\text{C.1})$$

Combining the Ramsey rule and the fact that $L_x = r/\xi$ (see Eq. (A.12)), we can derive

$$L_x = \beta(\rho - 1) \frac{\frac{L}{N} - L_f + \frac{1}{\beta(\rho - 1)} \frac{\theta}{\xi}}{(\rho - 1) \frac{\varepsilon - \beta}{\varepsilon - 1} + 1} . \quad (\text{C.2})$$

Finally, we derive labor used in the consumption goods sector as

$$L_C = (1 - \beta) \bar{N} \frac{\varepsilon}{\varepsilon - 1} (\rho - 1) \frac{\frac{L}{N} - L_f + \frac{1}{\beta(\rho - 1)} \frac{\theta}{\xi}}{(\rho - 1) \frac{\varepsilon - \beta}{\varepsilon - 1} + 1} . \quad (\text{C.3})$$

Taking derivatives yields

$$\frac{\partial L_x}{\partial \bar{N}} < 0; \quad \frac{\partial L_C}{\partial \bar{N}} < 0; \quad \frac{\partial g}{\partial \bar{N}} < 0. \quad (\text{C.4})$$

The optimum is found by putting dU_0/dN equal to zero (see Eq. (20)) which results in

$$N^{\text{SB}} = \frac{(\beta(\rho - 1)(\sigma - 1) - 1)L}{[(1 - \beta) + \beta\sigma] \left[L_f(\rho - 1) - \frac{\theta}{\xi\beta} \right]}. \quad (\text{C.5})$$

Substituting this solution for the second-best number of firms into Eq. (C.1) yields the solution for the second-best rate of growth

$$g^{\text{SB}} = \frac{\beta(\varepsilon - 1)\rho\xi L_f - \theta[\beta\rho(\sigma - 1)(\varepsilon - \beta) - \beta(\sigma - 1)(1 - \beta) - (1 - \beta)]}{\beta[\beta(\rho - 1)(\sigma - 1) - 1][(\varepsilon - \beta)(\rho - 1) + (\varepsilon - 1)]}. \quad (\text{C.6})$$

Finally, we compare the growth rate in the second-best social optimum with the first-best rate of growth (Eq. (15)). This yields

$$g^{\text{FB}} - g^{\text{SB}} = \frac{(1 - \beta)[\beta\xi L_f(\rho - 1) - \theta]}{\beta[\beta(\rho - 1)(\sigma - 1) - 1][(\varepsilon - \beta)(\rho - 1) + (\varepsilon - 1)]}. \quad (\text{C.7})$$

Using the conditions $\beta(\rho - 1) > \theta/(\xi L_f)$ and $\beta(\rho - 1) > 1/(\sigma - 1)$, it is easily seen that the growth rate in the SBSO is always lower than in the FBSO.

Acknowledgements

This paper is largely based on research done when both authors were affiliated to CentER, Tilburg University. Comments by two anonymous referees of this journal are gratefully acknowledged. We also would like to thank Erik Canton, Ben Heijdra, Theo van de Klundert, Ton van Schaik and Sjak Smulders for useful comments and stimulating discussion on earlier versions of this paper. Of course, the usual disclaimer applies.

References

Barro, R. J., and Sala-i-Martin, X. (1995): *Economic Growth*. New York: McGraw-Hill.

- Benassy, J. P. (1996): "Taste for Variety and Optimum Production Patterns in Monopolistic Competition." *Economics Letters* 52: 41–47.
- Benassy, J. P. (1998): "Is There Always too Little Research in Endogenous Growth with Expanding Product Variety?" *European Economic Review* 42: 61–69.
- Broer, D. P., and Heijdra, B. J. (2001): "The Investment Tax Credit under Monopolistic Competition." *Oxford Economic Papers* 53: 318–351.
- Ciccone, A., and Hall, R. E. (1996): "Productivity and the Density of Economic Activity." *American Economic Review* 86: 54–70.
- De Groot, H. L. F., and Nahuis, R. (1998): "Taste for Diversity and the Optimality of Economic Growth." *Economics Letters* 58: 291–295.
- De Groot, H. L. F., and Nahuis, R. (2000): "Note accompanying 'Optimal Product Variety and Economic Growth'." Download from <http://www.econ.vu.nl/re/mdw/henri>.
- Dixit, A. K., and Stiglitz, J. E. (1977): "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67: 297–308.
- Dosi, G. (1988): "Sources, Procedures and Microeconomic Effects of Innovation." *Journal of Economic Literature* 26: 1120–1171.
- Ethier, W. J. (1982): "National and International Returns to Scale in the Modern Theory of International Trade." *American Economic Review* 72: 389–405.
- Evans, G. W., Honkapohja, S., and Romer, P. (1998): "Growth Cycles." *American Economic Review* 88: 495–515.
- Grossman, G. M., and Helpman, E. (1991): *Innovation and Growth in the Global Economy*. Cambridge, Mass.: MIT Press.
- Hall, R. E. (1988): "Intertemporal Substitution in Consumption." *Journal of Political Economy* 96: 339–357.
- Holtz-Eakin, D., and Lovely, M. A. (1996): "Technological Linkages, Market Structure, and Optimum Production Policies." *Journal of Public Economics* 61: 73–86.
- Jovanovic, B. (1997): "Learning and Growth." In *Advances in Economics and Econometrics: Theory and Applications*, edited by D. M. Kreps and K. F. Wallis, Cambridge, Mass.: Cambridge University Press, pp. 318–339.
- Mankiw, N. G., and Whinston, M. D. (1986): "Free Entry and Social Inefficiency." *Rand Journal of Economics* 17: 48–58.
- Peretto, P. F. (1996): "Sunk Costs, Market Structure and Growth." *International Economic Review* 37: 895–923.
- Peretto, P. F., and Smulders, S. (1998): "Specialization, Knowledge Dilution, and Scale Effects in an IO-Based Growth Model." Tilburg: CentER Discussion Paper No. 9802.
- Roeger, W. (1995): "Can Imperfect Competition Explain the Difference between Primal and Dual Productivity Measures? Estimates for U.S. Manufacturing." *Journal of Political Economy* 103: 316–330.
- Romer, P. (1990): "Endogenous Technological Change." *Journal of Political Economy* 98: 71–102.
- Smulders, S. (1994): *Growth, Market Structure and the Environment. Essays on the Theory of Endogenous Economic Growth*. Tilburg: Ph.D. Thesis, Tilburg University.

- Smulders, S., and Van de Klundert, T. (1995): "Imperfect Competition and Growth with Firm-Specific R&D." *European Economic Review* 39: 139–160.
- Spence, M. (1976): "Product Selection, Fixed Costs, and Monopolistic Competition." *Review of Economic Studies* 43: 217–235.
- Van de Klundert, T., and Smulders, S. (1997): "Growth, Competition and Welfare." *Scandinavian Journal of Economics* 99: 99–118.
- Young, A. (1998): "Growth Without Scale Effects." *Journal of Political Economy* 106: 41–63.

Addresses of authors: – Henri L. F. de Groot, Vrije Universiteit, Department of Spatial Economics, De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands (e-mail: hgroot@econ.vu.nl); – Richard Nahuis, CPB Netherlands Bureau for Economic Policy Analysis and KUN Nijmegen p/a CPB Netherlands Bureau for Economic Policy Analysis, P.O. Box 80510, 2508 GM, The Hague, The Netherlands (e-mail: r.nahuis@cpb.nl)