

Optimal Fiscal Policy with Endogenous Product Variety *

Sanjay K. Chugh[†] Fabio Ghironi[‡]
University of Maryland Boston College and NBER

April 28, 2009

Abstract

We study optimal fiscal policy in an economy in which the number of product varieties is endogenous and in which sunk costs of development make products long-lived assets for the firms that sell them. Two main results emerge. First, depending on the particular form of variety aggregation, new product development may be either subsidized or taxed in the long run. This subsidization or taxation occurs on the dividends that firms pay out to their owners. In the most empirically relevant and intuitively appealing form of variety aggregation, however, the model delivers a clear prediction: dividend income should be taxed quite heavily. Because dividend income is a form of capital income, the result can be regarded as one of positive capital taxation, and the basic reason is the long-lived nature of product varieties. Second, regardless of the form of variety aggregation, in the short-run — i.e., in the face of business-cycle shocks — maintaining low volatility of labor income tax rates is *not* optimal. The standard deviation of the optimal labor income tax rate is over 2 percentage points, over an order of magnitude larger than benchmark tax-*smoothing* results in the Ramsey literature. Like long-run dividend taxation, the lack of tax smoothing is due to the long-lived nature of products in the economy. The higher is the rate of product turnover, the lower is the degree of optimal tax volatility. In the limiting case of purely static product creation decisions, tax smoothing is optimal.

*

[†]Email: chughs@econ.umd.edu.

[‡]Email: fabio.ghironi@bc.edu.

Contents

1	Introduction	4
2	Model	7
2.1	Households	7
2.2	Firms	9
2.2.1	Creation of New Varieties	10
2.2.2	Optimal Pricing	11
2.2.3	Production and Choice of Inputs	12
2.3	Government	12
2.4	Resource Constraint	12
2.5	Private-Sector Equilibrium	13
3	The Love of Variety Effect and Monopoly Power	14
4	Efficient Allocations and Intertemporal Efficiency	15
5	Ramsey Problem	18
6	Completeness of the Tax System	20
7	Optimal Policy: Analytical Results	22
8	Optimal Policy: Quantitative Results	25
8.1	Parameterization and Solution Strategy	26
8.2	Long-Run Optimal Policy	28
8.2.1	Dixit-Stiglitz Aggregation	28
8.2.2	Translog Aggregation	29
8.2.3	Benassy Aggregation	32
8.3	Short-Run Optimal Policy	35
8.3.1	Optimal Tax Rate Volatility	35
8.3.2	Dynamics of Real Variables	42
9	Robustness and Further Quantitative Experiments	43
9.1	Immediate Delivery to Market	43
9.2	Product Creation Subsidies	46
10	Conclusion	48

A	Derivation of Pricing Equation	52
B	Private-Sector Equilibrium	54
B.1	Compact Representation of Equilibrium	56
B.2	Aggregate Resource Constraint	56
C	Efficient Allocations	58
D	Derivation of Present-Value Implementability Constraint	60
D.1	Baseline Model	60
D.2	Immediate Delivery to Market	62
E	Optimal Long-Run Dividend Tax	63
F	Long-Run Elasticity of New Products with Respect to Taxes	65

1 Introduction

A large and resurgent literature has documented the importance of product creation and turnover for business cycle dynamics. Thus far, there has been little work on developing new insights for optimal macroeconomic policy based on this evidence. This paper is an early step towards that goal. In particular, we study the long-run and short-run properties of optimal fiscal policy in an economy in which product varieties are endogenous and firms engage in forward-looking product creation decisions based on the prospect of earning long-lived streams of monopoly profits. Taken together, these ideas seem to capture the microeconomic underpinnings of product turnover.

Our analysis yields two main results, one regarding the long run and one regarding the short run. In the long run, the model predicts that optimal dividend income taxation can be zero, positive, or negative. However, in the most empirically relevant and intuitively appealing version of the model, the dividend income tax rate should be positive in the long run and quite high — 50 percent, if we take the model literally. Dividend taxation in this case discourages an inefficiently high quantity of product development. In the short run, the model predicts that the optimal labor income tax rate is extremely volatile in the face of business cycle magnitude shocks. At a cyclical standard deviation of over 2 percentage points (around a long-run mean of roughly 20 percent), the labor income tax rate is over an order of magnitude more volatile than is optimal in benchmark Ramsey models. The model thus predicts that tax smoothing, the typical prescription in the basic Ramsey literature, is *not* optimal. The high volatility of labor tax rates is due to the persistence of product varieties in the economy: the more rapid is the rate of product turnover, the closer the model comes to prescribing the standard tax-smoothing result.

The environment in which the analysis is conducted is the model developed by Bilbiie, Ghironi, and Melitz (2007), who use an endogenous growth framework along the lines of Romer (1990) to study the business cycle implications of an endogenous, time-varying, and long-lived stock of product varieties. Their model is motivated by a long history of empirical research on the cyclical implications of product development and turnover, a research question that has recently received renewed attention due to the availability of highly-detailed micro-level data sets.¹ The Bilbiie, Ghironi, and Melitz (2007) framework — hereafter, BGM — makes many empirically-relevant predictions regarding business cycle dynamics, including its ability to match well the cyclical behavior of profits, the number of product varieties, and goods-market markups. For this latter aspect of their results, the crucial element of the BGM model is a translog aggregator over differentiated

¹A couple of representative papers in this most recent burst of work on this issue are Bernard, Redding, and Schott (2006) and Broda and Weinstein (2007). We certainly cannot do justice to the most recent activity in the literature, nor to the rich older literature on this topic, so we refer interested readers to the more complete list of references in Bilbiie, Ghironi, and Melitz (2007).

varieties, first developed by Feenstra (2003), that has the intuitively appealing property that an increase in the number of varieties available in the economy is associated with an increase in the degree of substitutability between any given pair of varieties. This aspect of aggregation is absent in the most commonly used specification of the Dixit-Stiglitz aggregator, which assumes a constant elasticity of substitution across varieties even if the number of varieties is endogenous.

The variety aggregator matters for the optimal long-run dividend income tax. If variety aggregation is translog (the aggregation that BGM shows has the most empirically attractive cyclical properties), the optimal long-run dividend tax is 50 percent. We prove this result in the full Ramsey model. The intuition for the result is that with translog aggregation and *zero* dividend taxation, the monopoly incentives governing product development are stronger than the beneficial effects of increased product variety on welfare. Too many products are thus developed in equilibrium. A dividend income tax, which effectively taxes monopoly profits — we assume firms pay out all their flow profits as dividends — corrects this distortion.² If variety aggregation is instead of the commonly-used Dixit-Stiglitz form, we prove that the long-run dividend tax is exactly zero. In the Dixit-Stiglitz case, the product-development incentive of profits and the variety effect exactly balance each other. We also consider the generalization of Dixit-Stiglitz aggregation proposed by Benassy (1996); in this aggregation, optimal dividend income taxes can be either positive or negative, depending on which of the two effects is stronger. Taken together, our results suggest that the optimal dividend income tax in the long run is *not* likely to be zero, and, based on other attractive properties of the model, most likely to be positive.

Our study is a classic Ramsey taxation analysis, featuring exogenous government spending that must be financed using proportional taxation and without reliance on lump-sum taxes. Underlying any Ramsey analysis is a pure social planning problem. For the BGM environment, the pure social planning solutions and the corrective (Pigovian) taxes needed to support those allocations were developed in detail by Bilbiie, Ghironi, and Melitz (2008). Their results provide the analytical basis for the results we obtain regarding long-run (Ramsey-) optimal taxation. Of particular importance for our work here is that Bilbiie, Ghironi, and Melitz (2008) — hereafter, BGM2 — determined the constellations of conditions for the markup incentives governing product development and the conditions governing the effects on welfare of product variety that are important for efficiency. It is the tradeoff of these two forces that shapes the long-run optimal dividend income tax.

Interestingly, the public revenue potential of dividend income taxes is not at all exploited by the Ramsey government in the long run. Instead, the long-run dividend income tax is driven *only* by the efficiency concerns identified by BGM2; their results carry over to a full Ramsey analysis. However,

²The precise 50-percent dividend tax rate follows from the functional form of the translog expenditure function; we provide details below.

we extend this aspect of BGM2 by reframing the characterization of efficiency in terms of marginal rates of substitution and appropriately-defined corresponding marginal rates of transformation, which is helpful for understanding the Ramsey policy using the key idea that optimal policy is all about creating optimal *wedges*. We also connect the optimal policy predictions to the classic results on capital income taxation dating back to Chamley (1986) and Judd (1985) as well as to more recent developments in the Ramsey literature on asset taxation.

Regarding the stochastic business cycle dynamics of optimal policy, an issue not studied by BGM2, our main result is that labor income tax rates are very volatile. In our baseline calibration, which is adopted from BGM, the cyclical standard deviation of the optimal labor income tax rate is over 2 percent, around a mean of about 20 percent. Thus, the classic tax-*smoothing* prescription that is the core prediction of many Ramsey analyses is not optimal here. Tax rates are over an order of magnitude more volatile than benchmark results in the Ramsey literature, and this volatility is independent of the form of variety aggregation. The feature of the environment most crucial for this short-run prediction is the long-lived nature of products. We show that a very high rate of product turnover reinstates the classic tax-smoothing result. At well over 50 percent per quarter, however, the rate of product turnover required for tax smoothing to be optimal is clearly counterfactually high. As the rate of product turnover approaches 100 percent, a *static* model of product creation and turnover, similar in idea to that studied in Jaimovich and Floetotto (2008), emerges. Our results show that the optimality of tax smoothing is not disrupted by product creation and turnover *per se*, but rather by product creation decisions that are inherently *dynamic* ones.³

More broadly, our work also connects to a recent branch of the optimal-policy literature that has been studying policy insights in a rich class of dynamic stochastic general equilibrium (DSGE) environments. Some examples include the monetary policy studies in frictional labor markets by Faia (2008), Thomas (2007), and Arseneau and Chugh (2008), the study of labor-income taxation in frictional labor markets by Arseneau and Chugh (2009), and the study of capital-income taxation by Arseneau, Chugh, and Kurmann (2008) in an environment featuring frictional markets for physical capital. The common idea tying together these studies and our work in this paper is the presence of forward-looking private-sector behavior in aspects richer than captured in standard RBC-based or standard New Keynesian-based models. This literature has shown that forward-looking behavior, apart from standard forward-looking capital accumulation and pricing behavior, can deliver new

³We also note that tax volatility in our model is not driven by any incompleteness of government debt markets, which is a well-understood point in Ramsey models since Aiyagari, Marcet, Marimon, and Sargent (2002). Thus, the only point of departure from a neoclassically-based complete-markets Ramsey setup, such as the textbook Chari and Kehoe (1999) presentation, is that product creation is subject to sunk upfront costs of development and governed by the promise of a long-lived stream of expost monopoly profits.

insights for some classic questions on optimal policy.

We emphasize that our results have nothing to do with incompleteness of the tax system, which can be an important issue in Ramsey models of taxation. As Chari and Kehoe (1999, p. 1679-1680) explain, an incomplete tax system is in place if, for at least one pair of goods in the economy, the government has *no* policy instrument that drives a wedge between the marginal rate of substitution (MRS) of those goods and the corresponding marginal rate of transformation (MRT). Well understood in the Ramsey literature is that incompleteness of tax systems can cause all sorts of “strange” policy prescriptions to arise. None of our taxation results arises due to any inability on the part of the government to create wedges between one or more MRS/MRT pairs. Indeed, in Section 6, we prove that our model features a complete set of policy instruments by demonstrating that the government has one unique policy tool for each wedge between MRS and MRT that it might want to create.

The rest of the paper is organized as follows. Section 2 describes the economic environment. Section 3 provides an intuitive discussion of the two key economic forces in the environment: the love of variety effect and the monopoly incentives for product development. Section 4 reframes the efficiency results of BGM2 in terms of marginal rates of substitution and marginal rates of transformation. Section 5 describes the Ramsey problem. Section 6 formally establishes a point we noted above, that the tax system in our analysis is complete. Section 7 builds on BGM2 and proves the Ramsey-optimal long-run dividend income tax. Section 8 uses a parameterized version of the model to characterize both long-run and short-run properties of optimal policy, including demonstrating the optimality of tax volatility. In this section, we conduct a variety of experiments that reveal that the most important feature governing the volatility of the tax rate is the rate of product turnover. Section 9 conducts a few robustness exercises. Section 10 concludes. Proofs and most algebraic derivations are relegated to the appendices.

2 Model

We set up the economic environment, describing in turn the behavior of households, the nature of product varieties and firms, the activities of the government, and the private-sector equilibrium.

2.1 Households

The representative household chooses state-contingent processes for the consumption index c_t and total hours worked h_t to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, h_t) \tag{1}$$

subject to a sequence of flow budget constraints

$$c_t + v_t x_{t+1} (N_t + N_{E,t}) + b_{t+1} = (1 - \tau_t^h) w_t h_t + R_t b_t + (v_t + (1 - \tau_t^d) d_t) x_t N_t, \quad (2)$$

in which x_{t+1} denotes the number of shares of a mutual fund (stock) the household carries between period t and $t + 1$, and v_t is its real (in units of the numeraire c_t) price. Each unit of the mutual fund held at the start of period t entitles the household to the receipt of a per-unit dividend d_t . In equilibrium, this dividend will be the flow profits from a given product line in the economy, of which there is a measure N_t in period t . The rest of the notation is standard: h_t denotes time spent working (thus, with time per period normalized to one, $1 - h_t$ is time spent in leisure), w_t is the market real wage, b_t is holdings of a real state-contingent government bond at the start of period t , and R_t is the realized gross interest rate on bond holdings carried into period t . Labor income earnings are taxed at the rate τ_t^h , and dividend income earnings are taxed at the rate τ_t^d .

The household purchases goods in and sells its labor on spot markets. The standard consumption-leisure optimality condition thus describes household optimal choice along this margin,

$$-\frac{u_h(c_t, h_t)}{u_c(c_t, h_t)} = (1 - \tau_t^h) w_t. \quad (3)$$

A standard bond demand function

$$u_c(c_t, h_t) = \beta E_t \{u_c(c_{t+1}, h_{t+1}) R_{t+1}\} \quad (4)$$

emerges from intertemporal optimization, as does the stock demand function

$$v_t = (1 - \delta) E_t \left\{ \frac{\beta u_c(c_{t+1}, h_{t+1})}{u_c(c_t, h_t)} (v_{t+1} + (1 - \tau_{t+1}^d) d_{t+1}) \right\}. \quad (5)$$

The stock demand function has the standard interpretation that the per-share price in any period is equal to the present-discounted value of future after-tax dividend flows.

Having optimally chosen the consumption index c_t , the household then chooses the quantity of each (symmetric) variety in order to minimize the total cost of purchasing the index c_t . With a symmetric and homothetic aggregator over varieties, the demand function for each available variety i is

$$c_{it} = \frac{\partial P_t}{\partial \tilde{P}_{it}} c_t, \quad (6)$$

where P_t is the nominal price of the consumption index (imagine simply multiplying the budget constraint above by P_t) and $\tilde{P}_{it}$ is the nominal price of symmetric variety i . Below, we will define $\rho_t \equiv \tilde{P}_{it}/P_t$ as the relative price of (symmetric) variety i in terms of the consumption index. We also discuss below the specifications we use for the variety aggregator.

2.2 Firms

There is a continuum of identical firms that produce and sell output, so we can restrict attention to a representative firm. The representative firm is modeled as being a “large firm” that produces “many” varieties. Decisions about product creation in various product lines are assumed to be completely independent of each other.⁴ This large firm formulation facilitates aggregation and, with the assumption of independence across varieties, yields the same equilibrium conditions as the formulations of BGM and BGM2, which technically do not employ the large-firm approach.⁵

Expressed in nominal terms, the intertemporal profit maximization problem of the representative (large) firm is

$$\max_{\tilde{P}_t, N_{t+1}, N_{E,t}} E_0 \sum_{t=0}^{\infty} \Xi_{t|0} \left[(1 - \tau_t^d)(\tilde{P}_t - P_t mc_t) N_t q(\tilde{P}_t) - P_t mc_t f_{E,t} N_{E,t} \right]. \quad (7)$$

The discount factor the firm applies to its profits has two components. The first component, which applies to total period- t profits, is the standard intertemporal marginal rate of substitution (MRS) of households (who are the owners of firms). Between date zero and date t , this MRS is denoted $\Xi_{t|0}$.⁶ The second component of the discount factor used by the firm applies only to operating profits from existing varieties in production — that is, to the dividend payments the firm makes to its owners. The firm discounts the dividends it disburses in period t by the after-tax rate $(1 - \tau_t^d)$. In equilibrium, these dividends are, in nominal terms and on a per-product basis, $P_t d_t = (\tilde{P}_t - P_t mc_t) q(\tilde{P}_t)$. Because there is no principal-agent problem between firms and households (the owners), firms discount flows of profits at the same intertemporal rate at which households discount future dividend receipts, and this properly takes into account the tax rate τ_t^d .

The profit function is written in such a way that it anticipates the equilibrium in which all (ex-ante differentiated) goods are symmetric.⁷ The rest of the notation is as follows. The pre-existing stock of product varieties in period t is N_t . Sales of each of these varieties occurs in a

⁴This independence assumption regarding development decisions for various product lines implies that the firm does not behave strategically across product lines in its product introduction decisions. Put another way, the firm does not “internalize” the effects of entering one product market on its other product lines. Such an assumption can be justified by supposing that separate “brand managers” oversee each product line and that they communicate little with each other, or more explicitly, are actually encouraged to *compete* with each other.

⁵One can view our large-firm approach as analogous to many recent general-equilibrium macro models that feature search and matching frictions in various markets. In such models, the “large firm” assumption (for example, a representative firm that has to search individually for the “many” employees it seeks to hire) also facilitates aggregation and ignores strategic considerations.

⁶In equilibrium, we will have as usual that $\Xi_{t|0} = \frac{\beta^t u_{ct}}{u_{c0}}$.

⁷We can express the profit function this way because we analyze only an equilibrium symmetric across differentiated varieties. Apriori, the profit function would be $\sum_{t=0}^{\infty} \Xi_{t|0} \left[(1 - \tau_t^d) \int_0^{N_t} (\tilde{P}_{it} - P_t mc_t) q(\tilde{P}_{it}) di - P_t mc_t f_{E,t} N_{E,t} \right]$, where i indexes the potentially asymmetric varieties, of which there exists a measure N_t in period t .

monopolistically-competitive goods market. The term mc_t is the real marginal cost of producing a (symmetric) variety in terms of the overall consumption basket/index; we link mc_t to the production function below. The demand function $q(\tilde{P}_t)$ for a given variety is taken as given by the firm.⁸ As described above, $\tilde{P}_t$ is the nominal price of a symmetric variety, and P_t is the nominal price of the consumption index. Finally, in each period, the firm chooses how many new varieties $N_{E,t}$ to develop and attempt to bring to market in the subsequent period. Creating a new product entails a fixed cost $f_{E,t}$ that is denominated in effective labor units.

We analyze the problem of the firm in three steps: we first characterize its decisions regarding new product creation, we then describe the firm's optimal pricing function, and we finally describe the production process and characterize the firm's choice of inputs.

2.2.1 Creation of New Varieties

The representative firm takes as constraint the law of motion for the total number of varieties it produces and sells,

$$N_{t+1} = (1 - \delta)(N_t + N_{E,t}). \quad (8)$$

The exogenous fraction δ of varieties existing in period t disappear from the market before period $t + 1$. Of the newly developed products $N_{E,t}$ in period t , the fraction δ also never make it to market in period $t + 1$. Although we do not endogenize the product destruction margin, this captures in a simple aggregative way the idea of product life cycles.

Letting λ_t denote the Lagrange multiplier on the period- t law of motion for the number of varieties, optimization of (7) (taking (8) as constraint) with respect to $N_{E,t}$ and N_{t+1} yields, respectively, the first-order conditions

$$-P_t mc_t f_{E,t} + \lambda_t(1 - \delta) = 0 \quad (9)$$

and

$$-\lambda_t + E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d) \left(\tilde{P}_{t+1} - P_{t+1} mc_{t+1} \right) q(\tilde{P}_{t+1}) + (1 - \delta) \lambda_{t+1} \right] \right\}. \quad (10)$$

Eliminating the multiplier λ_t between these expressions gives

$$mc_t f_{E,t} = (1 - \delta) E_t \left\{ \Xi_{t+1|t} \frac{P_{t+1}}{P_t} \left[(1 - \tau_{t+1}^d) \left(\frac{\tilde{P}_{t+1}}{P_{t+1}} - mc_{t+1} \right) q(\tilde{P}_{t+1}) + mc_{t+1} f_{E,t+1} \right] \right\}, \quad (11)$$

⁸Note that we write $q(\tilde{P})$ to stand for the demand function, rather than $c(\tilde{P})$; this is because we also will have (exogenous) government consumption. To keep things symmetric, we suppose that government purchases are also of an “index good,” and then that the bundle of differentiated varieties the government consumes is identical to the bundle consumed by private households. Hence $q(\tilde{P})$ subsumes both private and public demand for a (symmetric) variety. The algebraic units of q is physical units of differentiated variety.

which is a free-entry condition in the market for any given variety. We refer to this condition as the product-creation condition. Because in our analysis of optimal policy, we want to abstract from monetary issues in order to focus on fiscal policy, we normalize $P_{t+1} = P_t \forall t$.⁹ Finally, defining $\rho_t \equiv \tilde{P}_t/P_t$ as the relative price of a symmetric variety (in terms of the consumption index) and denoting $q(\tilde{P}_t)$ simply by q_t , the product-creation condition can be expressed as

$$mc_t f_{E,t} = (1 - \delta) E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d) (\rho_{t+1} - mc_{t+1}) q_{t+1} + mc_{t+1} f_{E,t+1} \right] \right\}. \quad (12)$$

Comparing the product-creation condition (12) with the stock-pricing condition (5) derived above, it is clear that in equilibrium $v_t = mc_t f_{E,t}$ and $d_t = (\rho_t - mc_t) q_t$.

2.2.2 Optimal Pricing

Based on the constrained optimization problem described above and given a number of product varieties, the first-order condition with respect to the nominal price of a symmetric variety is

$$N_t q(\tilde{P}_t) + \tilde{P}_t N_t q'(\tilde{P}_t) - P_t mc_t N_t q'(\tilde{P}_t) = 0. \quad (13)$$

Defining $\zeta_t \equiv \frac{\tilde{P}_t}{q(\tilde{P}_t)} q'(\tilde{P}_t)$ as the price elasticity of demand for a symmetric variety, we have

$$\frac{\tilde{P}_t}{P_t} \left(1 + \frac{1}{\zeta_t} \right) = mc_t \Rightarrow \frac{\tilde{P}_t}{P_t} = \left(\frac{\zeta_t}{1 + \zeta_t} \right) mc_t. \quad (14)$$

The profit-maximizing relative price is thus in general an endogenously-time-varying markup over real marginal cost. Denoting by ρ_t the optimal relative price of a symmetric variety, $\rho_t \equiv \tilde{P}_t/P_t$, and by μ_t the gross markup, $\mu_t \equiv \frac{\zeta_t}{1 + \zeta_t}$, the optimal pricing rule can be expressed more compactly as

$$\rho_t = \mu_t mc_t. \quad (15)$$

As we discuss further below, both the markup μ and the relative price ρ are in general functions of the number of available varieties, N . The precise relationship depends on the specific aggregator over varieties.

A more complete derivation of the optimal pricing rule is provided in Appendix A. Looking ahead to the Ramsey problem, though, an important feature of this pricing rule is that markups in goods markets, μ_t , may fluctuate over the business cycle and, separately, relative prices of varieties

⁹Because we assume that the government issues fully state-contingent debt, this assumption is without loss of generality. More precisely, as is well understood from explicitly monetary Ramsey analyses (for example, see the overview in Chari and Kehoe (1999)), the basic reason for variations in the aggregate price level under the optimal policy is precisely to *achieve* state contingency in real government debt payments. If the government issues state contingent debt to begin with, there is no Ramsey rationale for state-contingent movements in the aggregate price level. Thus, we shut this channel down without loss of generality.

ρ_t may fluctuate over the business cycle. By construction, neither type of fluctuation occurs in standard monopolistically-competitive DSGE models based on the Dixit-Stiglitz framework and a fixed number of varieties, in which markups and relative prices are constant over the business cycle.¹⁰ To the extent that optimal policy is concerned with stabilizing relative prices, fluctuations in μ_t (or, equivalently, fluctuations in mc_t and/or fluctuations in ρ_t) have the potential to be a key driver of our model's cyclical optimal-policy prescriptions.

2.2.3 Production and Choice of Inputs

Given a pricing rule and a pre-existing stock of varieties, the firm produces a quantity of goods of each variety to satisfy demand at its chosen price. Production of each existing variety takes place using a linear-in-labor technology. Letting h_t^G denote labor used to produce q_t units of a particular variety, the existing-goods-producing technology is $q_t = z_t h_t^G$, where z_t is an exogenous level of productivity that is common across varieties. In equilibrium, symmetry of costs across varieties implies symmetry of prices and production across varieties.

The technology involved in setting up a new product variety is also linear. Letting h_t^E denote the *exogenous* quantity of labor used to produce one new variety, $z_t h_t^E$ is the production function for a new variety. Linking this to the fixed cost of entry $f_{E,t}$ defined above, we have simply that $f_{E,t} = z_t h_t^E$. Because z_t and h_t^E are both exogenous, the fixed cost $f_{E,t}$ is obviously also exogenous.

2.3 Government

The government must finance an exogenous stream of government purchases $\{g_t\}_{t=0}^\infty$ via proportional labor income taxation, dividend income taxation, and one-period state-contingent debt. Its flow budget constraint is

$$\tau_t^h w_t h_t + \tau_t^d d_t x_t N_t + b_{t+1} = g_t + R_t b_t, \quad (16)$$

2.4 Resource Constraint

We can write the resource frontier as

$$c_t + g_t + \rho_t N_{E,t} f_{E,t} = \rho_t z_t h_t, \quad (17)$$

in which, by labor market clearing, $h_t = h_t^G N_t + h_t^E N_{E,t}$. We obtain the resource frontier by summing the flow household budget constraint (2) and the flow government budget constraint (16) and then making several substitutions using other first-order conditions of the economy; a full

¹⁰Nominal rigidities in a standard New Keynesian model of course disrupt unit relative prices in goods markets, and hence cause markups to fluctuate over the business cycle. It is this departure of relative prices from unity that drive, among things, optimal-policy prescriptions in such models.

derivation is presented in Appendix B. A crucial feature to note about the resource frontier is that ρ_t , which in the decentralized economy represents a relative price, appears. We discuss this “technological” aspect of ρ in Section 3.

2.5 Private-Sector Equilibrium

We study only equilibria that are symmetric across product varieties — that is, we concentrate on equilibria in which production and sales of an identical quantity q occurs of each of the N varieties. This is justified by the assumption is symmetry of costs across varieties. Now that we are at the stage of constructing the equilibrium, we make explicit the equilibrium dependence of the markup and the relative price of a given product on the total stock of products in the economy — thus, we now explicitly write $\mu(N_t)$ and $\rho(N_t)$. The analytic forms of these functions are presented in Section 3.

As shown in detail in Appendix B, the definition of a (symmetric) private-sector equilibrium can be expressed quite compactly. Specifically, an equilibrium can be defined as a set of endogenous processes, $\{c_t, h_t, N_{t+1}, N_{E,t}, b_t\}_{t=0}^{\infty}$, for given exogenous processes $\{z_t, g_t, \tau_t^h, \tau_t^d, f_{E,t}\}_{t=0}^{\infty}$, that satisfy five sequences of conditions: a static consumption-leisure equilibrium condition

$$-\frac{u_{ht}}{u_{ct}} = \frac{(1 - \tau_t^h)}{\mu(N_t)} z_t \rho(N_t); \quad (18)$$

an intertemporal product creation condition

$$\rho(N_t) f_{E,t} = (1 - \delta) E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d) \mu(N_t) \left(1 - \frac{1}{\mu(N_{t+1})} \right) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \frac{\mu(N_t)}{\mu(N_{t+1})} \rho(N_{t+1}) f_{E,t+1} \right] \right\}; \quad (19)$$

the law of motion for the number of product varieties

$$N_{t+1} = (1 - \delta)(N_t + N_{E,t}); \quad (20)$$

the aggregate resource constraint

$$c_t + g_t + \rho(N_t) N_{E,t} f_{E,t} = \rho(N_t) z_t h_t; \quad (21)$$

and the flow government budget constraint

$$\tau_t^h w_t h_t + \tau_t^d d_t N_t + b_t = g_t + R_t b_{t-1}. \quad (22)$$

Substituting the equilibrium expressions for the real wage and dividend payments, $w_t = z_t m c_t = \frac{z_t \rho(N_t)}{\mu(N_t)}$ and $d_t = (\rho(N_t) - m c_t) \left(\frac{c_t + g_t}{N_t \rho(N_t)} \right)$, into the latter casts the government budget constraint also in terms of only the endogenous processes listed above.

We now turn to an intuitive discussion of the importance of $\rho(N)$ and $\mu(N)$ in this framework and describe the functional forms for aggregators.

Dixit-Stiglitz	Benassy	Translog
$\mu(N_t) = \mu = \frac{\theta}{\theta-1}$	$\mu(N_t) = \mu = \frac{\theta}{\theta-1}$	$\mu(N_t) = 1 + \frac{1}{\sigma \tilde{N}_t}$
$\rho(N_t) = N_t^{\mu_t-1} = N_t^{\frac{1}{\theta-1}}$	$\rho(N_t) = N_t^\kappa$	$\rho(N_t) = \exp\left(-\frac{1}{2} \frac{\tilde{N}_t - N_t}{\sigma \tilde{N}_t N_t}\right)$, $\tilde{N} \equiv \text{Mass}(\text{potential products})$
$\epsilon(N_t) = \mu - 1$	$\epsilon(N_t) = \kappa$	$\epsilon(N_t) = \frac{1}{2\sigma \tilde{N}_t} = \frac{1}{2}(\mu(N_t) - 1)$

Table 1: For each of the three variety aggregators, the markup function and love of variety function.

3 The Love of Variety Effect and Monopoly Power

In the subsequent two sections, we describe efficiency and the Ramsey problem. Before these formal presentations, however, it is useful to briefly discuss intuitively some of the key efficiency insights for this environment developed in BGM2. The two economic forces that are pitted against each other along the product creation dimension are the beneficial effects of variety on welfare and the monopoly signals that provide ex-ante incentives for the creation of varieties. In an arbitrary equilibrium, these two forces trade off in ways that depend on the particular form of variety aggregation.

Following BGM and BGM2, we consider three specifications for variety aggregation: the familiar Dixit-Stiglitz aggregator, the generalization explained by Benassy (1996), and the translog aggregator of Feenstra (2003).¹¹ Table 1 presents a summary of markups $\mu(N)$ and relative prices $\rho(N)$, derived in BGM2, for each of the three variety aggregators. More about the functional forms is presented when we discuss quantitative issues in Section 8. For now, what is important is that each aggregator implies, for a given stock N of product varieties, a different monopolistic markup and a different “love of variety effect” — that is, the effect of variety on welfare.

The chance to earn a long-lived stream of monopoly markups $\mu(N)$ provides the ex-ante incentives for firms to pay the sunk costs of product development. Ex-post, monopoly rents impose a distortion on the economy. There is thus a societal cost, apart from the physical costs of product creation, of entry by firms into a new product line.

On the other hand, an expansion of the set of products available in the economy is socially desirable provided consumers’ preferences exhibit “love of variety.” For a given consumption index c and given nominal expenditures Pc , welfare is higher if those expenditures are spread over a larger number of product varieties; this is the essence of love-of-variety preferences, as is well-known in

¹¹As Benassy (1996) himself makes clear, this generalization actually appeared in a working paper version of Dixit and Stiglitz (1977), although not in the published version. Because Benassy (1996) re-discovered this result, we will refer to this aggregator as the “Benassy specification.”

the trade literature.¹² This love of variety effect is summarized by the relationship $\rho(N)$. In a decentralized equilibrium, $\rho(N)$ also turns out to be the relative price of a (symmetric) variety. Thus, $\rho(N)$ plays a dual role in the environment: not only is it a price in the decentralized economy, but it is also a fundamental part of the allocation from the point of view of a social planner and, hence, from the point of view of the Ramsey government. As Table 1 shows, $\rho'(N) > 0$ in all three cases. An expansion in the number of varieties is thus welfare-improving, *ceteris paribus*.

The love of variety effect and the monopoly signals governing product creation thus define the tension shaping the efficient stock of products. How $\mu(N)$ compares to $\rho(N)$ is a basic theme of the efficiency and optimal-policy results described below. Rather than working directly with the function $\rho(N)$, however, for much of the analysis, it is convenient to characterize the love of variety effect in elasticity form by defining

$$\epsilon(N_t) \equiv \rho'(N_t) \frac{N_t}{\rho(N_t)}. \quad (23)$$

Regardless of whether things are framed in terms of $\rho(N)$ or $\epsilon(N)$, a social planner, and hence also the Ramsey government, takes into account the allocational consequences of the number of product varieties.

4 Efficient Allocations and Intertemporal Efficiency

Before presenting and analyzing the Ramsey problem, we list the conditions that characterize the allocations $\{c_t, h_t, N_{t+1}, N_{E,t}\}$ that a social planner would choose. The following discussion of efficiency is helpful for understanding the Ramsey government's optimal policy.

Social efficiency is characterized by the four (sequences of) conditions

$$-\frac{u_{ht}}{u_{ct}} = z_t \rho(N_t), \quad (24)$$

$$\rho(N_t) f_{E,t} = (1 - \delta) E_t \left\{ \Xi_{t+1|t} \left[\epsilon(N_{t+1}) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \rho(N_{t+1}) f_{E,t+1} \right] \right\}, \quad (25)$$

$$N_{t+1} = (1 - \delta)(N_t + N_{E,t}), \quad (26)$$

and

$$c_t + g_t + \rho(N_t) N_{E,t} f_{E,t} = \rho(N_t) z_t h_t, \quad (27)$$

where, as above, $\Xi_{t+1|t} \equiv \frac{\beta u_{ct+1}}{u_{ct}}$. These conditions pin down efficient paths for the set of processes $\{c_t, h_t, N_{t+1}, N_{E,t}\}_{t=0}^{\infty}$. As derived in Appendix C, the efficiency conditions (24) and (25) are

¹²See, for example, Grossman and Helpman (1991). In the trade literature, the “love-of-variety” effect is also often interpreted as returns-to-specialization. This alternative interpretation is based on assuming that production of final goods depends on using differentiated input goods; we instead have assumed this differentiation directly in consumers’ utility aggregator. The idea is clearly the same.

obtained by maximizing social welfare, given by (1), subject to the technological frontier defined by (26) and (27). In these efficiency conditions, we emphasize the fact that a planner takes into account the love of variety effect on welfare by writing $\rho(N_t)$ — in the decentralized economy, this effect works through (and is identical to) the relative price of a variety, as noted in Section 3. Comparing (24) and (25) with their decentralized equilibrium counterparts (18) and (19) is instructive. The subsequent observations regarding attainment of efficiency in the decentralized economy were previously established in BGM2.

Comparison of the equilibrium condition (18) with the efficiency condition (24) shows that setting $\tau_t^h = 1 - \mu(N_t)$ achieves efficiency along the static consumption-leisure margin. With monopoly power, the gross markup is $\mu(N_t) > 1$. Achieving static efficiency thus requires a *subsidy* to labor income, a point well-understood in models with monopoly power and as shown by BGM2 for this environment. With positive government spending that must be financed via non-lump-sum taxation, however, the average labor income tax rate will be *positive*. The Ramsey equilibrium will thus feature a distorted consumption-leisure margin, a standard feature of Ramsey equilibria with endogenous labor supply.

Comparison of the equilibrium condition (19) with the intertemporal efficiency condition (25) shows it is less obvious how, or even whether, efficiency can be achieved along the product creation margin along a dynamic stochastic Ramsey equilibrium path. It is easier to make statements here regarding the long-run, deterministic, Ramsey equilibrium. In the deterministic steady-state, intertemporal efficiency is described by

$$\frac{1}{\beta} = (1 - \delta) \left[\frac{\epsilon(N) \left(\frac{c+g}{N} \right) + \rho(N) f_E}{\rho(N) f_E} \right], \quad (28)$$

which is obtained by imposing steady-state on (25). We have written this condition in a way that lines up with the concepts presented in Proposition 1 below. For a given level of aggregate demand of the final index $c + g$, this condition fully characterizes the long-run efficient number of varieties N .

For the subsequent analysis of the Ramsey problem and solution, it is useful to restate the conditions describing efficient allocations in terms of marginal rates of substitution (MRS) and corresponding marginal rates of transformation (MRT), now in fully dynamic (but deterministic) form.

Proposition 1. *The MRS and MRT for the pairs (c_t, h_t) , and (c_t, c_{t+1}) are defined by*

$$\begin{aligned} MRS_{c_t, h_t} &\equiv -\frac{u_{ht}}{u_{ct}} & MRT_{c_t, h_t} &\equiv z_t \rho(N_t) \\ MRS_{c_t, c_{t+1}} &\equiv \frac{u_{ct}}{\beta u_{ct+1}} & MRT_{c_t, c_{t+1}} &\equiv (1 - \delta) \left[\frac{\epsilon(N_{t+1}) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \rho(N_{t+1}) f_{E, t+1}}{\rho(N_t) f_{E, t}} \right]. \end{aligned}$$

Proof. Follows from the efficiency conditions (24) and (25). $\square$

Each MRS in Proposition 1 of course has the standard interpretation as a ratio of relevant marginal utilities. Similarly, each MRT has the interpretation as a ratio of the marginal products of an appropriately-defined transformation frontier.¹³ The static MRT in the first line above shows that the variety effect, embodied in $\rho(N_t)$, can be thought of as part of the description of the *technology* of the economy. Indeed, this can also be concluded from the fact that $\rho(N_t)$ appears in the resource frontier, (27), of the economy.

The intertemporal MRT (IMRT) in the second line above deserves further discussion. This IMRT can be understood based on the primitives of the environment. In the standard one-sector RBC model, in which there is only one type of produced good, it is straightforward to define the IMRT using the economy-wide intertemporal resource constraint. Due to our model's two-sector structure (the goods-producing sector and the varieties-producing sector), however, defining the IMRT (in terms of final consumption goods) is slightly more involved but instructive to understand intuitively.

By definition, the IMRT (in terms of final consumption goods) measures how many additional units of c_{t+1} the economy can achieve if one unit of c_t is foregone, holding constant output of all other goods in the economy. If c_t is reduced by one unit, resources are freed up to produce additional varieties in period t , $N_{E,t}$. From the resource frontier (27), it is clear that a unit reduction in c_t means new varieties in period t can be increased by $\frac{1}{\rho(N_t)f_{E,t}}$, holding constant total output $\rho(N_t)z_t h_t$. The law of motion of product varieties, given by (26), shows that a fraction δ of these new varieties disappear before reaching the production stage in period $t+1$. In period $t+1$, there are thus $(1-\delta)\frac{1}{\rho(N_t)f_{E,t}}$ new products that will actually be produced.

In period $t+1$, these $(1-\delta)\frac{1}{\rho(N_t)f_{E,t}}$ additional varieties are transformable into consumption c_{t+1} through two channels. First, they can be directly transformed back into consumption: the period- $t+1$ resource frontier shows that this transformation from varieties back to consumption happens at the rate $\rho(N_{t+1})f_{E,t+1}$ (simply the inverse of the transformation that occurred in period t , except measured according to the period $t+1$ state). Second, the additional stock of product varieties yields a “dividend” in terms of extra production. Because we consider only allocations that are symmetric across existing varieties (that is, production of each of the N existing varieties at any date is identical), this additional production in physical terms is $\frac{c_{t+1}+g_{t+1}}{N_{t+1}}$. Due to the love of variety effect, output of this additional variety is scaled by the variety effect $\epsilon(N_{t+1})$. Hence, the total extra c_{t+1} the economy can achieve by reducing c_t by one unit is given by

$$(1-\delta) \left[\frac{\epsilon(N_{t+1}) \left(\frac{c_{t+1}+g_{t+1}}{N_{t+1}} \right) + \rho(N_{t+1})f_{E,t+1}}{\rho(N_t)f_{E,t}} \right].$$

¹³We have in mind a very general notion of transformation frontier as in Mas-Colell et al (1995, p. 129).

The following corollary characterizes the efficient allocations in terms of MRSs and MRTs.

Corollary 1. *The solution to the social planner's problem is characterized by the law of motion for product varieties (26), the goods resource constraint (27), along with*

$$\frac{MRS_{c_t, h_t}}{MRT_{c_t, h_t}} \equiv \frac{-u_{h_t}/u_{c_t}}{z_t \rho(N_t)} = 1, \quad (29)$$

and

$$\frac{MRS_{c_t, c_{t+1}}}{MRT_{c_t, c_{t+1}}} \equiv \frac{u_{c_t} \rho(N_t) f_{E,t}}{\beta u_{c_{t+1}} (1 - \delta) \left[\epsilon(N_{t+1}) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \rho(N_{t+1}) f_{E,t+1} \right]} = 1. \quad (30)$$

Proof. Obvious from Proposition 1 and the efficiency conditions (24)-(25) □

Corollary 1 shows that the efficient allocations in our model can be described in terms of “zero-wedge conditions” between MRSs and MRTs. This way of understanding efficiency is of course standard, but given the novelty of the model, it is important to show how to precisely express efficiency in terms of the zero-wedge expressions (29) and (30). Conditions (29) and (30) recast the efficiency results of BGM2 in a way that makes more transparent the Ramsey analysis to follow. This is especially important because, as Chari and Kehoe (1999, p. 1674) emphasize, optimal tax theory is really about the determination of optimal wedges between MRSs and MRTs. In what follows, we take expressions (29) and (30) as the conditions that define zero wedges.

5 Ramsey Problem

We discuss in some detail the formulation of the Ramsey problem because a novel issue arises in its construction relative to the existing Ramsey literature. Stated in the most straightforward way, the problem of the Ramsey government is to raise revenue to finance exogenous government expenditure through labor income taxes, dividend income taxes, and issuance of one-period real state-contingent debt in such a way that maximizes the welfare of the representative household, subject to the equilibrium conditions of the economy. Assuming, as we do, that the Ramsey government can commit to state-contingent policy functions, this straightforward statement of the Ramsey problem translates as choosing state-contingent functions for τ^h , τ^d , government debt returns, and all the endogenous equilibrium objects listed in Section 2.5 in order to maximize (1) subject to (18)-(22) (with the markup and variety conditions appropriate for the particular aggregator specification adopted). Such a formulation of the Ramsey problem is known in the literature as the dual formulation. The dual formulation retains price and policy variables as explicit choice objects of the Ramsey government.

A commonly used representation of the Ramsey problem, the primal formulation, instead eliminates all price and policy variables using the economy's equilibrium conditions. Doing so casts

the Ramsey problem in terms of directly choosing allocations subject to feasibility and a small set of conditions that describe implementation as a decentralized equilibrium. In a second step, one then residually constructs via equilibrium conditions the prices and policies as the hyperplanes that support the Ramsey allocation. In our environment, the pure primal approach does not fully characterize the Ramsey equilibrium.

We instead study the Ramsey problem using a hybrid between the dual approach and the primal approach. Our approach is informed by the fact that the sequence of asset prices $\{v_t\}$ from date $t = 0$ onwards plays a central role in shaping allocations in this environment — in particular, product-creation decisions — and hence in shaping optimal policy. A central goal of the Ramsey government in this environment is thus to affect the allocative asset prices, which it can do by appropriately choosing dividend tax rates $\{\tau_t^d\}$. As we discuss next, this requires leaving the tax rates τ_t^d as a direct choice variable in the Ramsey problem. This sort of point, that of taxation being used to affect asset prices in the Ramsey equilibrium, was first noted by Armenter (2008) and further developed by Arseneau, Chugh, and Kurmann (2008). It is an issue that does not arise in Ramsey models featuring capital accumulation modeled in the standard one-sector RBC setup because the price of capital is by construction always unity and thus cannot be affected by taxes.

As is standard in the Ramsey literature, one component of our hybrid formulation is thus a summary of the sequence of government budget constraints, the sequence of consumption-leisure optimality conditions, and the sequence of asset (stock) demand conditions in a single present-value implementability condition (PVIC). The present-value representation is available to us because debt markets are assumed to be complete. As shown in Appendix D, the PVIC for this economy is

$$E_0 \sum_{t=0}^{\infty} \beta^t [u_{ct}c_t + u_{ht}h_t] = u_{c0} \left[R_b b_0 + (v_0 + (1 - \tau_0^d)d_0)N_0 \right]. \quad (31)$$

This recasting eliminates τ^h and the state-contingent bond returns from the list of directly-chosen Ramsey objects because they can both be constructed residually from the equilibrium conditions given a Ramsey allocation.

The Ramsey problem is now reduced to a choice of state-contingent functions for $\{c_t, h_t, N_{t+1}, N_{E,t}, \tau_t^d\}$ in order to maximize (1) subject to the PVIC (31), the sequence of product creation conditions (19), the sequence of laws of motion (20), and the sequence of resource constraints (21). Note that neither (19) nor (20) is captured by the PVIC. What makes this formulation not a pure primal Ramsey problem is the presence of the sequence of τ^d appearing in the Ramsey choice set and in the sequence of product creation conditions. Thus, we simply compute a first-order condition of this Ramsey problem with respect to τ_t^d . The fact that we optimize directly with respect to a policy variable is what sets our formulation apart from a strict primal one.

A few further issues arise, which are standard in complete-markets Ramsey models featuring

asset taxation. The first issue regards the deterministic Ramsey steady-state. In the long run of a complete-markets Ramsey model, the government's bond position is indeterminate. This is a well-known issue in Ramsey models; hence, the steady-state level of bonds is simply chosen as a parameter. The second issue regards Ramsey dynamics. Because we assume government debt markets are complete *and* because time- t dividend income taxes are akin to a lump-sum, state-contingent levy on the pre-determined stock of product varieties in the economy, it is impossible to pin down uniquely, along the business cycle (i.e., in the presence of shocks), a path of state-contingent dividend income taxes *and* a path of state-contingent government bond returns. We resolve this indeterminacy by always fixing the dividend income tax rate to its long-run value (which *is* pinned down) and allowing bond returns to fluctuate.

The Ramsey problem features forward-looking constraints in the sequence of product creation conditions and laws of motion for varieties. We focus on policy from a timeless perspective and ignore the effects of transitions from arbitrary initial conditions to the asymptotic steady state. Thus, the Ramsey planner at time zero respects the time $t = -1$ forward-looking product creation condition and law of motion for the number of varieties, which requires appending appropriate time $t = -1$ Lagrange multipliers to the Ramsey problem.¹⁴ Throughout, the first-order conditions of the Ramsey problem are assumed to be necessary and sufficient and that all allocations are interior. Finally, as is common in Ramsey taxation problems, we assume full commitment. Thus, we emphasize that none of our results is driven by the use of a discretionary policy.

6 Completeness of the Tax System

As we discussed in the introduction, an important issue in models of optimal taxation is whether or not the assumed tax instruments constitute a *complete* tax system.¹⁵ In this section, we establish that the tax system is complete in our model. Establishing this is important for two reasons. First, at a technical level, proving completeness reaffirms that the Ramsey problem as formulated in Section 5 is indeed the correct Ramsey problem. As shown by Chari and Kehoe (1999, p. 1680), Correia (1996), Armenter (2008), and many others, incompleteness of the tax system typically requires imposing additional constraints that reflect the incompleteness. Incompleteness is not

¹⁴Thus, as in Khan, King, and Wolman (2003), Faia (2008), Arseneau and Chugh (2008), and others, the initial state of the economy is assumed to be the asymptotic Ramsey steady state. Doing this achieves analysis of policy dynamics from what is commonly referred to as the “timeless” perspective, which captures the idea that the optimal policy has already been operational for a long time.

¹⁵For convenience, we restate the definition of Chari and Kehoe (1999, 1679-1680) that an *incomplete* tax system is in place if, for at least one pair of goods in the economy, the government has *no* policy instrument that drives a wedge between the marginal rate of substitution (MRS) and the corresponding marginal rate of transformation (MRT). If this is not the case, then the tax system is instead said to be *complete*.

an issue in our model, therefore we do not need to impose additional constraints. Second, it is well-known in Ramsey theory that incomplete tax systems can lead to a wide range of “non-standard” policy prescriptions in which some instruments stand in for the ability to create certain wedges that *cannot*, by assumption of the available tax instruments, be created in a decentralized economy. Proving completeness therefore establishes that none of our results is due to any policy instrument serving as imperfect proxies for other, unavailable, instruments.

As we showed in Section 4, there are two independent MRS/MRT pairs in the environment, and expressions (29)-(30) state the efficiency conditions in terms of these MRS/MRT pairs. Completeness of the tax system requires that each of these margins is affected by (at least) one policy instrument. To establish completeness, we first express explicitly in terms of MRS/MRT pairs the private-sector equilibrium conditions that are the analogs of the efficiency conditions (29)-(30).

Using (18), (19), and the definitions of MRSs and MRTs presented in Proposition 1, we have that in the decentralized economy

$$\frac{MRS_{c_t, h_t}}{MRT_{c_t, h_t}} = \frac{1 - \tau_t^h}{\mu(N_t)} \quad (32)$$

and

$$\frac{MRS_{c_t, c_{t+1}}}{MRT_{c_t, c_{t+1}}} = \frac{(1 - \tau_{t+1}^d)\mu(N_t) \left(1 - \frac{1}{\mu(N_{t+1})}\right) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}}\right) + \frac{\mu(N_t)}{\mu(N_{t+1})}\rho(N_{t+1})f_{E, t+1}}{\epsilon(N_{t+1}) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}}\right) + \rho(N_{t+1})f_{E, t+1}}. \quad (33)$$

We have expressed (33) in deterministic form because our analytical results regarding the intertemporal margin are only for the deterministic Ramsey steady state.

Next, we express in the same way the first-order conditions of the Ramsey problem (which are derived in Appendix E); doing so gives

$$\frac{MRS_{c_t, h_t}}{MRT_{c_t, h_t}} = 1 + \xi \left\{ 1 + \frac{u_{cct}c_t}{u_{ct}} + \frac{1}{z_t\rho(N_t)} \left[\frac{u_{hht}h_t}{u_{ct}} + \frac{u_{ht}}{u_{ct}} \right] \right\}, \quad (34)$$

and

$$\frac{MRS_{c_t, c_{t+1}}}{MRT_{c_t, c_{t+1}}} = 1, \quad (35)$$

where ξ is the Lagrange multiplier associated with the PVIC in the Ramsey problem.

We are now ready to establish completeness of the tax system.

Proposition 2. *Completeness of the Tax System.*

1. *Suppose that the Ramsey allocation converges to a deterministic steady state. In the steady state of a decentralized equilibrium, the two policy instruments τ^h and τ^d can uniquely create, in the margins defined by conditions (29) and (30), the wedges implied by the steady-state Ramsey allocation.*

2. *Along the dynamic path of any decentralized equilibrium, the instrument τ_t^h can uniquely create, in the margin defined by condition (29), the wedge implied by the dynamic path of the Ramsey allocation.*

Proof. For the first part of the statement, compare the steady-state version of (33) with that of (35). The second part of the statement follows directly because (32) and (34) are both static conditions. $\square$

Note that, with reference to wedges between the IMRS and IMRT, Proposition 2 covers only the steady state. As is well-understood in Ramsey theory and as discussed in Section 5, outside of steady state, if both a state-contingent dividend-income tax (which is a type of capital-income tax) and state-contingent government debt returns are possible, both cannot be simultaneously pinned down.¹⁶ In fact, as in virtually all of the literature, our analytical results regarding dividend income taxes (which are akin to capital income taxes) are only steady-state results (while our results regarding the optimal labor income tax cover both in- and out-of-steady-state allocations).

Having proven completeness of the tax system, we next proceed to characterizing the optimal policy by solving the Ramsey problem as defined in Section 5 and decentralizing it using policy instruments. We emphasize that none of the policy prescriptions that emerge from our analysis is because one (or more) of the policy instruments that is assumed available is acting as an imperfect substitute for a policy instrument that is assumed unavailable.

7 Optimal Policy: Analytical Results

Our first main result is that in the long-run Ramsey equilibrium, efficiency is achieved along the product-creation margin, and the long-run dividend income tax that supports this outcome depends on the particular form of variety aggregation. This result is stated in the following Proposition:

Proposition 3. *Long-Run Optimal Dividend Income Tax.* *In the deterministic steady state of the Ramsey equilibrium, the optimal dividend income tax rate is characterized by*

$$1 - \tau^d = \frac{\epsilon(N)}{\mu(N) - 1}, \quad (36)$$

and this tax supports the socially-efficient level of product creation in the long run.

Proof. See Appendix E. $\square$

A striking aspect about the Ramsey-optimal long-run dividend income tax is that it is identical to the purely Pigovian tax derived by BGM2. In particular, the goal of dividend taxation is to align

¹⁶For further discussion, Chari and Kehoe (1999, p. 1708) or Ljungqvist and Sargent (2004, p. 500-502).

the beneficial effects of product variety with net monopoly markups. As (36) shows, this requires $\epsilon(N) = \mu(N) - 1$, just as BGM2 found. The analysis of BGM2 is about corrective (Pigovian) taxes because it abstracts from public-finance considerations by assuming the availability of lump-sum taxation. Proposition 3 shows that endogenizing public finance considerations does not alter at all their long-run taxation predictions.

The way to understand this result is quite simple and relies on the analogy that BGM and BGM2 both offer that the stock of product varieties is similar to the stock of physical capital in a neoclassical RBC model. In an RBC economy absent public finance considerations, efficiency of course entails zero intertemporal distortions in the long run. The decentralized equilibrium achieves exactly this outcome due to the perfectly-competitive nature of all markets in the RBC economy. When one considers optimal taxation in a full public finance analysis in the RBC economy, the optimal capital income tax is zero, and this policy supports intertemporal efficiency — this is the famous result of Chamley (1986) and Judd (1985).

BGM2 show that to achieve an efficient quantity of product varieties — which is a type of *intertemporal* efficiency due to the dynamic nature of product creation and turnover — corrective taxation may be needed depending on how the variety effect, summarized by $\epsilon(N)$, trades off against the markup signal $\mu(N) - 1$ in the decentralized equilibrium. Setting the long-run dividend tax according to (36) equalizes the welfare benefit of variety with the net markup, which supports efficient product accumulation.

Our full Ramsey analysis shows that the optimality of maintaining efficiency along the product accumulation margin is robust to public-finance considerations. In this sense, the result in Proposition 3 is a manifestation of the Chamley (1986) and Judd (1985) result applied to product accumulation. Thus, the analogy offered by BGM and BGM2 that the stock of varieties is akin to the stock of capital in an RBC economy is helpful not only for understanding business-cycle results (as in BGM), but also helpful for understanding optimal-taxation implications.

While the formal proof of Proposition 3 is left to Appendix E, the structure of the proof is simple. The condition characterizing the Ramsey-optimal intertemporal margin turns out to be, in the long run, identical to the long-run intertemporal efficiency condition (28). The proof of this is slightly more involved than the Pigovian proof of BGM2 because of the presence of the equilibrium conditions (19) and (31) as constraints on the Ramsey problem; but the upshot is that condition (28) nevertheless characterizes the Ramsey government's intertemporal allocation. Decentralizing the long-run Ramsey allocation then requires comparing (28) to the deterministic steady-state version of the private-sector equilibrium condition (19). The optimal long-run dividend tax in (36) then follows.

In addition to drawing an analogy with Chamley (1986) and Judd (1985), this result can also be

connected to Albanesi and Armenter (2007), who provide a unified framework with which to think about long-run (physical) capital taxation in a wide variety of public-finance environments. An important distinction they make is that long-run intertemporal efficiency may or may not require a long-run zero capital income tax. Specifically, their work highlights that the essence of the Chamley (1986) and Judd (1985) result is not that zero-capital-taxation *per se* is optimal, but rather that it is optimal because it supports a zero wedge between intertemporal marginal rates of substitution and intertemporal marginal rates of transformation. Whether a zero intertemporal wedge requires a zero capital tax or a nonzero capital tax then depends on the details of the economic environment. From a technical standpoint, the method of proof of Proposition 3 also fits squarely into the sufficient conditions for zero long-run intertemporal distortions provided by Albanesi and Armenter (2007).¹⁷

The detail of the economic environment that matters here for the precise τ^d that decentralizes the zero-intertemporal-wedge allocation is the form of variety aggregation. With the three variety aggregations considered, the optimal dividend income tax can be positive, negative, or zero. Specifically, based on the love-of-variety and markup functions $\epsilon(N)$ and $\mu(N)$ presented in Section 3, the optimal long-run dividend income tax rate in the Dixit-Stiglitz aggregation is

$$\tau_{DS}^d = 1 - \frac{\frac{\theta}{\theta-1} - 1}{\frac{\theta}{\theta-1} - 1} = 0; \quad (37)$$

in the translog aggregation is

$$\tau_{TRANSLOG}^d = 1 - \frac{\frac{1}{2\sigma N}}{\frac{1}{\sigma N}} = 0.5; \quad (38)$$

and in the Benassy aggregation is

$$\tau_{BENASSY}^d = 1 - \frac{\kappa}{\frac{\theta}{\theta-1} - 1} = 1 - \kappa(\theta - 1). \quad (39)$$

There is thus a tight analogy between optimal capital income taxation in the basic RBC model, as first derived by Chamley (1986) and Judd (1985), and the optimal dividend income tax in the presence of purposeful product creation and turnover. The Ramsey allocation features zero intertemporal distortions in the long run in order to support the efficient stock of product varieties, very much like the fact that the Chamley (1986) and Judd (1985) result is, at its core, about supporting the efficient stock of physical capital in the long run. With monopoly power and the love of variety effects competing against each other in the decentralized economy, however, supporting zero intertemporal distortions in the presence of product turnover may require zero, positive, or negative dividend income taxes, depending on the particular form of variety aggregation.

¹⁷This is because the sequence of product stocks does not appear in the sequences of *equilibrium* constraints faced by the Ramsey government and appears only in *technological* constraints, which is the key sufficiency condition in Albanesi and Armenter (2007).

Unless one believes literally in the translog aggregator, it is difficult to offer a precise numerical target for the long-run dividend income tax based on our analysis because there is little empirical evidence about the variety effect. This point is also noted by Benassy (1996), who calls for further research on the issue. To the extent that the Dixit-Stiglitz aggregation is a knife-edge parameterization, however, we think one fairly sound conclusion from Proposition 3 is that *zero* dividend income taxation is likely not the right policy prescription to offer. Based on the success BGM found in matching a host of business cycle properties with the translog aggregation, we lean towards that as the most favored model with which to also consider optimal taxation — in which case dividend income ought to be taxed fairly heavily (at 50 percent) in the long run. As mentioned in Section 3, the translog aggregation also has *a priori* appeal because it captures the idea that the larger the mix of available products, the closer substitutes they are, an idea articulated by neither the Dixit-Stiglitz nor the Benassy aggregations. Based on these grounds, our preferred reading of our results is that *positive* dividend income taxation is optimal in the long run.

We close this discussion by relating our result on dividend taxation to the analysis of the optimal quantity of R&D by Benassy (1998). Benassy (1998) applied his own (Benassy (1996)) variety aggregator to the Romer (1990) endogenous growth model to ask whether too much or too little R&D occurred in the decentralized economy relative to the social optimum. His answer was that it depended on whether the love of variety effect was stronger than or weaker than the markup effect, and he concluded that there was no basis for offering normative prescriptions due to the lack of empirical evidence about the variety effect. As we mentioned in the introduction, the BGM framework is essentially the Romer (1990) endogenous growth environment with growth stripped out. So, although Benassy (1998) does not go all the way to drawing policy prescriptions and does not consider a full DSGE Ramsey analysis, our long-run analysis can be viewed as a detrended version of his analysis. If we were limited to the Benassy aggregation in forming our conclusions, we agree with Benassy (1998) that there is no basis for recommending even a *sign* for the optimal dividend income tax because the sign of $\tau_{BENASSY}^d = 1 - \kappa(\theta - 1)$ depends on $\kappa(\theta - 1)$, which are both simply parameters.

8 Optimal Policy: Quantitative Results

We now turn to a numerical characterization of the long-run and short-run Ramsey equilibria. Before presenting results, we describe parameter settings and computational methodology.

8.1 Parameterization and Solution Strategy

As in BGM and BGM2, we consider three different aggregation functions of varieties in the consumption index: Dixit-Stiglitz aggregation, Benassy (1996)-type aggregation, and Feenstra (2003)-type translog aggregation.¹⁸ In the well-understood Dixit-Stiglitz case, the final consumption index c is composed of the underlying varieties c_i according to $c = \left[\int_0^N c_i^{(\theta-1)/\theta} di \right]^{\frac{\theta}{\theta-1}}$. Note the presence of the endogenous (and, in the dynamic model, time-varying) measure N of varieties. We set $\theta = 6$, which delivers a net markup of 20 percent, $\mu = \frac{\theta}{\theta-1}$, and a relative price of a symmetric variety of $\rho = N^{\mu-1}$.

In the Benassy case, the aggregator is the more general

$$c = N^{\kappa+1-\frac{\theta}{\theta-1}} \left[\int_0^N c_i^{(\theta-1)/\theta} di \right]^{\frac{\theta}{\theta-1}}. \quad (40)$$

Clearly, if $\kappa = \frac{\theta}{\theta-1} - 1$, the Dixit-Stiglitz specification is recovered. The advantage of the Benassy specification is that it disentangles the parameter governing the markup, θ , from the separate parameter $\kappa \geq 0$ governing the love of variety effect. As shown in Table 1, with Benassy aggregation, the markup of a symmetric variety continues to be given by $\mu = \frac{\theta}{\theta-1}$, but the relative price of a symmetric variety is given by $\rho = N^\kappa$. We continue using $\theta = 6$. With Benassy aggregation, we study the steady-state Ramsey policy for various values of κ . However, all the dynamic results presented hold κ fixed at $\frac{\theta}{\theta-1} - 1$; this is because, in terms of dynamics, the results are qualitatively the same as with alternative values of κ . It is only in the steady state that the more general Benassy aggregation yields qualitatively different results regarding optimal policy than Dixit-Stiglitz aggregation.

There is no closed-form direct specification for translog aggregation. Instead, the primitive is the expenditure function across varieties. Feenstra (2003) provides detailed analysis; here, we simply note that in the translog case, the markup is given by $\mu = 1 + \frac{1}{\sigma N}$ and the relative price of a symmetric variety is $\rho = \exp\left(-\frac{1}{2} \frac{\tilde{N}-N}{\sigma \tilde{N} N}\right)$, with the parameter $\tilde{N}$ interpreted as the potential total measure of varieties, the subset N of which actually exists and is produced. We set a very loose upper bound for the potential measure of goods, $\tilde{N} = 10^9$, which captures the idea that there is an unbounded number of goods in the potential product space. The parameter σ is set to $\sigma = 7.7$, which delivers an average net markup of 20 percent, matching the Dixit-Stiglitz and Benassy cases.

The final parameter regarding goods markets to be chosen is the rate of obsolescence of a given product variety. We adopt the calibration of BGM and set $\delta = 0.025$. Given the quarterly frequency of our model, this means that in the span of one year roughly 10 percent of products become obsolete and hence are no longer produced and sold. In the model, this probability of

¹⁸For more details than presented here, we refer interested readers to BGM and BGM2.

product obsolescence is independent of how long a product has been on the market. As the results of our experiments below show, δ is a key parameter for Ramsey dynamics.

For household preferences over leisure and the aggregate consumption index, we use the standard functional form

$$u(c, h) = \ln c + \zeta \ln(1 - h), \quad (41)$$

with ζ chosen so that steady-state hours worked in the Ramsey equilibrium featuring Dixit-Stiglitz aggregation of varieties is 0.25; the resulting value is $\zeta = 2.9$. We employ a quarterly calibration, hence the subjective discount factor is set to $\beta = 0.99$, which delivers an annual real interest rate of approximately four percent.

The exogenous productivity and government spending shocks follow AR(1) processes in logs,

$$\ln z_t = \rho_z \ln z_{t-1} + \epsilon_t^z, \quad (42)$$

$$\ln g_t = (1 - \rho_g) \ln \bar{g} + \rho_g \ln g_{t-1} + \epsilon_t^g, \quad (43)$$

where $\bar{g}$ denotes the steady-state level of government spending, which we calibrate in our baseline model with Dixit-Stiglitz aggregation to constitute 19 percent of steady-state output in the Ramsey allocation. The resulting value is $\bar{g} = 0.047$, which we hold constant across all experiments and all specifications of our model. The innovations ϵ_t^z and ϵ_t^g are distributed $N(0, \sigma_{\epsilon^z}^2)$ and $N(0, \sigma_{\epsilon^g}^2)$, respectively, and are independent of each other. We choose parameters $\rho_z = 0.95$, $\rho_g = 0.97$, $\sigma_{\epsilon^z} = 0.007$, and $\sigma_{\epsilon^g} = 0.02$, consistent with the RBC literature and Chari and Kehoe (1999). Also regarding policy, we set long-run government debt, a parameter in a complete-markets Ramsey model, so that the steady-state government debt-to-GDP ratio in the Dixit-Stiglitz case is 0.4, in line with long-run evidence for the U.S. economy and with the calibrations of Schmitt-Grohe and Uribe (2004a, 2004b), Chugh (2007), and Arseneau and Chugh (2008). Finally, the fixed cost of creating a new product is fixed at $f_E = 1$ and does not vary over the business cycle. Shutting down cyclical fluctuations in f_E achieves the greatest possible comparability with the basic Ramsey literature, in which it is almost always only exogenous variation in TFP and government purchases that drive the economy.

To numerically characterize the non-stochastic Ramsey steady-state, we apply a nonlinear numerical solver to the first-order conditions of the Ramsey problem with all exogenous processes set equal to their mean realizations. To study dynamics, we approximate our model by linearizing in levels the Ramsey first-order conditions for time $t > 0$ around the non-stochastic steady-state of these conditions. We use our approximated decision rules to simulate time-paths of the Ramsey equilibrium in the face of a complete set of TFP and government spending realizations, the shocks to which we draw according to the parameters of the laws of motion described above. We conduct 500 simulations, each 200 periods long. For each simulation, we compute first and second moments

and report the medians of these moments across the 500 simulations. Our numerical method is our own implementation of the perturbation algorithm described by Schmitt-Grohe and Uribe (2004c). As in Khan, King, and Wolman (2003) and others, we assume that the initial state of the economy is the asymptotic Ramsey steady state. As we mentioned above, we assume throughout, as is common in the literature, that the first-order conditions of the Ramsey problem are necessary and sufficient and that all allocations are interior. Finally, we again point out that we assume full commitment at date $t = 0$ by the Ramsey government to its policy function.

8.2 Long-Run Optimal Policy

We begin by presenting several aspects of the long-run Ramsey equilibrium. As stated in Section 2.5, we study only equilibria that are symmetric across varieties.

8.2.1 Dixit-Stiglitz Aggregation

As proven in Section 7, the long-run optimal tax on dividend income is zero if variety aggregation is Dixit-Stiglitz. Figure 1 confirms this numerically and also presents other aspects of the long-run Ramsey equilibrium as the elasticity parameter θ varies between $\theta = 3$ (which implies a net markup of 50 percent) and $\theta = 20$ (which implies a net markup of 5 percent). Many variables behave in quite predictable ways as θ increases (which means the markup decreases): the number of varieties decreases because smaller ex-post monopoly power provides weaker ex-ante incentive for product creation; the relative price of each variety decreases as the degree of competition increases; and the equilibrium quantity of labor rises as monopoly distortions, which spill over into the labor market, diminish. These results are in line with the conventional understanding of how the Dixit-Stiglitz monopolistically-competitive environment works.

More surprising is the result shown in the lower left panel of Figure 1 that total GDP declines as the markup declines. This result is the opposite of the long-run relationship between output and markups that arises in a standard Dixit-Stiglitz model in which the number of product varieties is fixed. The key to understanding the result with endogenous product variety is that creation of new varieties occurs in a separate sector than the sector in which production of consumption goods occurs. Virtually all macro-Dixit-Stiglitz models feature only one sector. In contrast, in this two-sector environment, the profit incentive to create new products diminishes as the markup shrinks, hence output of the product-development sector declines. Confirming this is the result, not shown in Figure 1, that steady-state N_E declines monotonically as the markup falls. Thus, the monotonic decline in *total* GDP as θ rises is mostly a reflection of the diminished output of this sector.

We say “mostly” because, as the lower middle panel shows, long-run consumption of the final

basket behaves non-monotonically with respect to θ . The non-monotonic long-run response of consumption is directly a consequence of the non-monotonicity of the long-run optimal labor income tax rate with respect to θ , which is shown in the upper middle panel. With our assumed utility form and under Dixit-Stiglitz aggregation, the steady-state consumption-leisure equilibrium condition can be expressed as

$$\frac{\zeta(1-h)}{c} = (1-\tau^h) \left(\frac{\theta-1}{\theta} \right) N^{\mu-1} \left(= \frac{(1-\tau^h)}{\mu(N)} \rho(N) \right). \quad (44)$$

With $N^{\mu-1}$ strictly decreasing in θ and h strictly increasing in θ , this expression confirms that the non-monotonicity in c is a direct reflection of the non-monotonicity in τ^h .

8.2.2 Translog Aggregation

Figure 2 presents the same long-run Ramsey responses as Figure 1 for the case of translog aggregation. In Figure 2, we vary the parameter σ between $\sigma = 1$ and $\sigma = 60$ (recall our baseline is $\sigma = 7.7$), which achieves the same variation in markups between 5 percent and 50 percent as in the experiments in Figure 1. All other parameters are held constant from the Dixit-Stiglitz economy.

With translog aggregation, a decline in the markup (shown in the upper right panel) is associated with a decline in the number of varieties (shown in the middle panel). Just as in the Dixit-Stiglitz case, this is because of the reduced profit incentive for product creation.¹⁹ The middle left panel shows that the decline in the number of varieties is associated in the Ramsey equilibrium with a rise in the relative price of a variety. To those familiar with BGM and BGM2, this association may be surprising. Indeed, simply examining the relation $\rho(N) = \exp\left(-\frac{1}{2} \frac{\tilde{N}-N}{\sigma \tilde{N} N}\right)$ for the translog case shows that, *ceteris paribus*, a decline in N should be associated with a decline in ρ (that is, $\rho'(N) > 0$). However, this is a comparative static on only N .

To understand what seems to be the opposite prediction here, we need to consider the *joint* long-run behavior of relative prices *and* the markup as σ varies. We noted above that for the translog case, the markup is given by $\mu = 1 + \frac{1}{\sigma N}$. Substituting $\frac{1}{\sigma N} = \mu - 1$ into the translog expression for the relative price, we have

$$\rho = \exp\left(-\frac{1}{2}(\mu-1)\left(1 - \frac{N}{\tilde{N}}\right)\right), \quad (45)$$

which shows that, unlike the Dixit-Stiglitz case, a squeezing of markups is associated with a *rise* in the relative price of a given variety. This then reconciles the result shown in the upper right panel of Figure 2 with that shown in the middle left panel.

¹⁹And, indeed, we once again confirm, although not shown in Figure 2, that the number of new products created, N_E , declines as the markup shrinks.

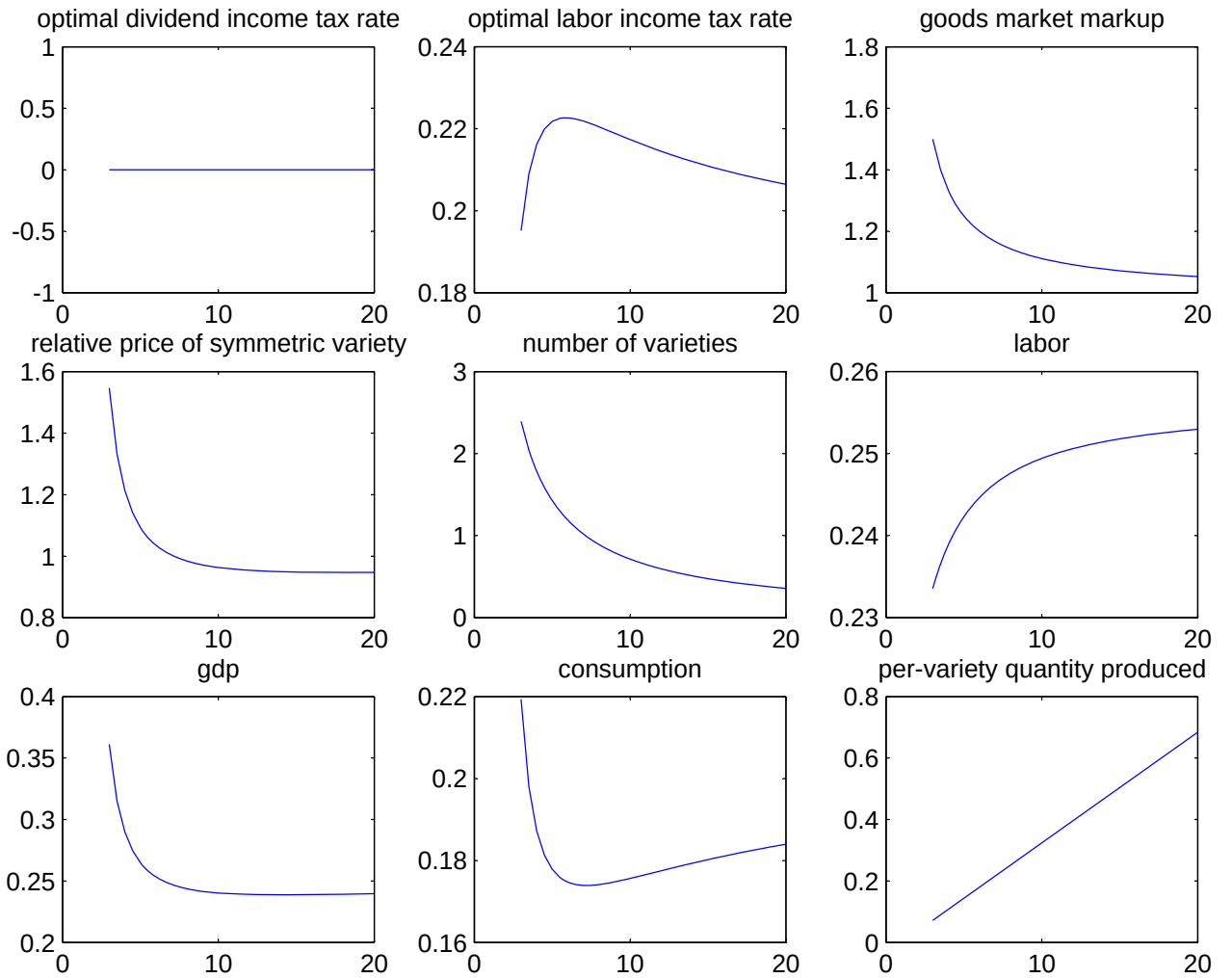


Figure 1: Long-run Ramsey equilibrium for Dixit-Stiglitz aggregation as elasticity parameter varies, $\theta \in [3, 20]$.

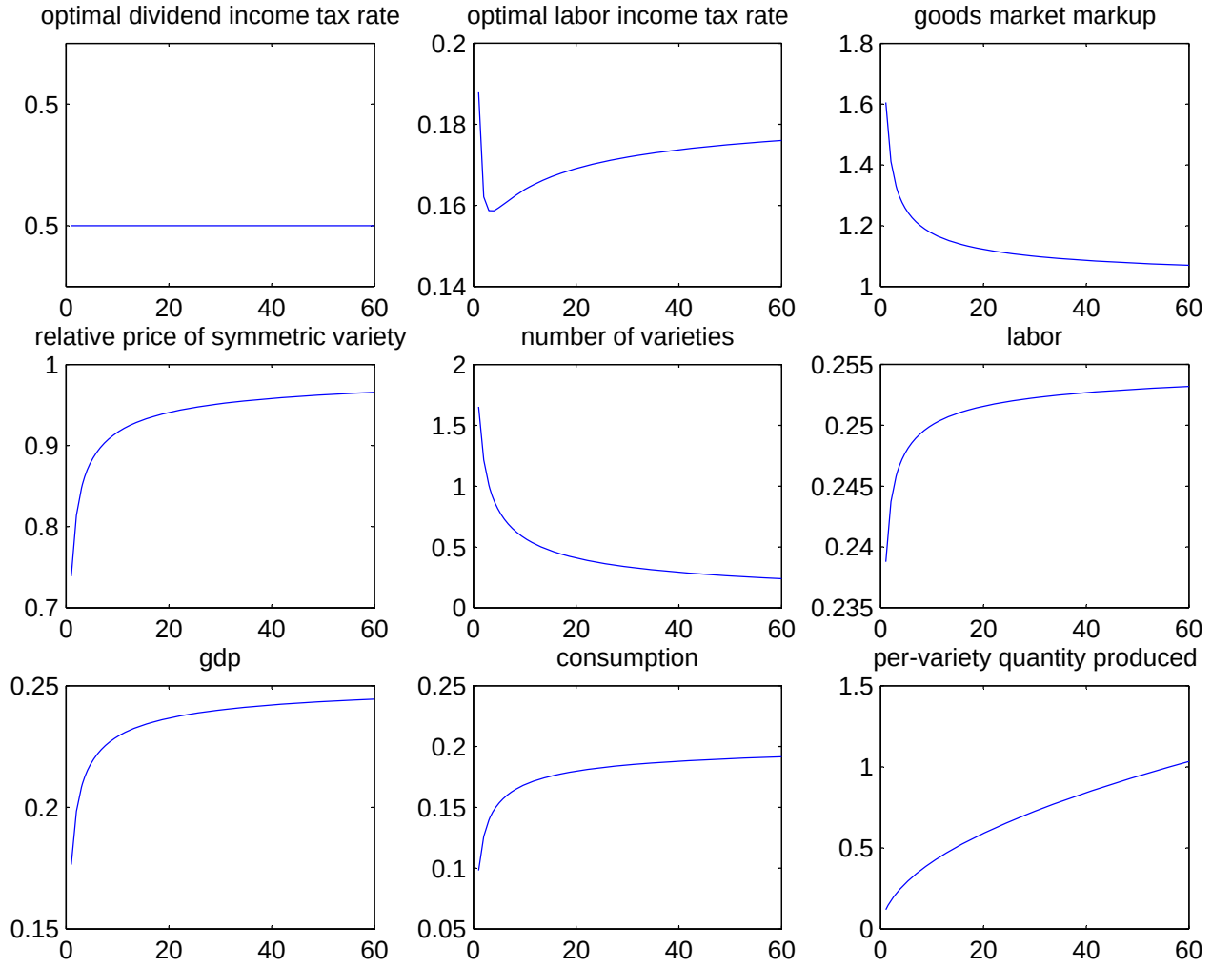


Figure 2: Long-run Ramsey equilibrium for translog aggregation as elasticity parameter varies, $\sigma \in [1, 60]$.

Regarding the long-run Ramsey policy, under translog aggregation, the dividend income tax rate is 50 percent, as proven in Section 7 and confirmed in the upper left panel of Figure 2. The labor income tax, with the exception of at extremely low values for σ , monotonically rises as the markup falls.

8.2.3 Benassy Aggregation

Finally, in Figure 3, we repeat the same experiments for the case of Benassy aggregation, which disentangles markups from love of variety. We keep the net markup fixed at 20 percent by holding $\theta = 6$, as in the Dixit-Stiglitz case, and vary the parameter κ to achieve variation in the real marginal cost of production between roughly 0.83 (which corresponds to the inverse of a gross markup of 1.2 in a standard Dixit-Stiglitz setting when the number of firms is fixed at measure one) and unity (which corresponds to the marginal cost of production in perfectly-competitive goods markets). This experiment requires varying κ between $\kappa = 0$ and $\kappa = 0.32$.²⁰ The Dixit-Stiglitz case emerges at $\kappa = \frac{\theta}{\theta-1} - 1 = 0.20$, in which case the variety effect on welfare and the profit incentive on entry are aligned, $\epsilon(N) = \mu(N) - 1$, without any government intervention.

As the upper left panel of Figure 3 shows, if $\kappa < 0.20$, a dividend income tax is optimal, whereas if $\kappa > 0.20$, a dividend income subsidy is optimal. This result distinguishes the Benassy case from the Dixit-Stiglitz and translog cases: the optimal dividend income tax rate varies across Ramsey equilibria, whereas it is always zero in the Dixit-Stiglitz case and always 50 percent in the translog case. The intuition is quite simple and follows from the result in Proposition 3: with $\kappa < \frac{\theta}{\theta-1} - 1$, the profit incentives provided by the markup are stronger than the variety effect on welfare. Under a zero dividend tax rate, the number of products in equilibrium is thus inefficiently high. A positive dividend income tax discourages some product creation. The effects work in reverse for the case of $\kappa > \frac{\theta}{\theta-1} - 1$.

It is interesting that a simple disentangling of the love of variety effect from monopoly power admits a wide range of optimal dividend-income tax policies. However, because it is difficult to determine which is the more empirically relevant case — $\kappa < \frac{\theta}{\theta-1} - 1$ or $\kappa > \frac{\theta}{\theta-1} - 1$ — these predictions unfortunately are not very strong. This point echoes Benassy's (1998) conclusion that, if Benassy aggregation is used, further empirical evidence on the love of variety effect is needed before useful normative statements can be offered. To the extent that pure Dixit-Stiglitz aggregation is a knife-edge case of Benassy aggregation, however, the optimality of a long-run *zero* dividend-income tax is clearly called into serious question by an overall reading of our results.

The other results in Figure 3 are not comparable to those in Figures 1 and 2 in a straightforward

²⁰Technically, the Ramsey equilibrium does not exist for $\kappa = 0$, so we actually set $\kappa = 0.001$ as a proxy. We explain in the next footnote why $\kappa = 0$ leads to a breakdown of the Ramsey equilibrium.

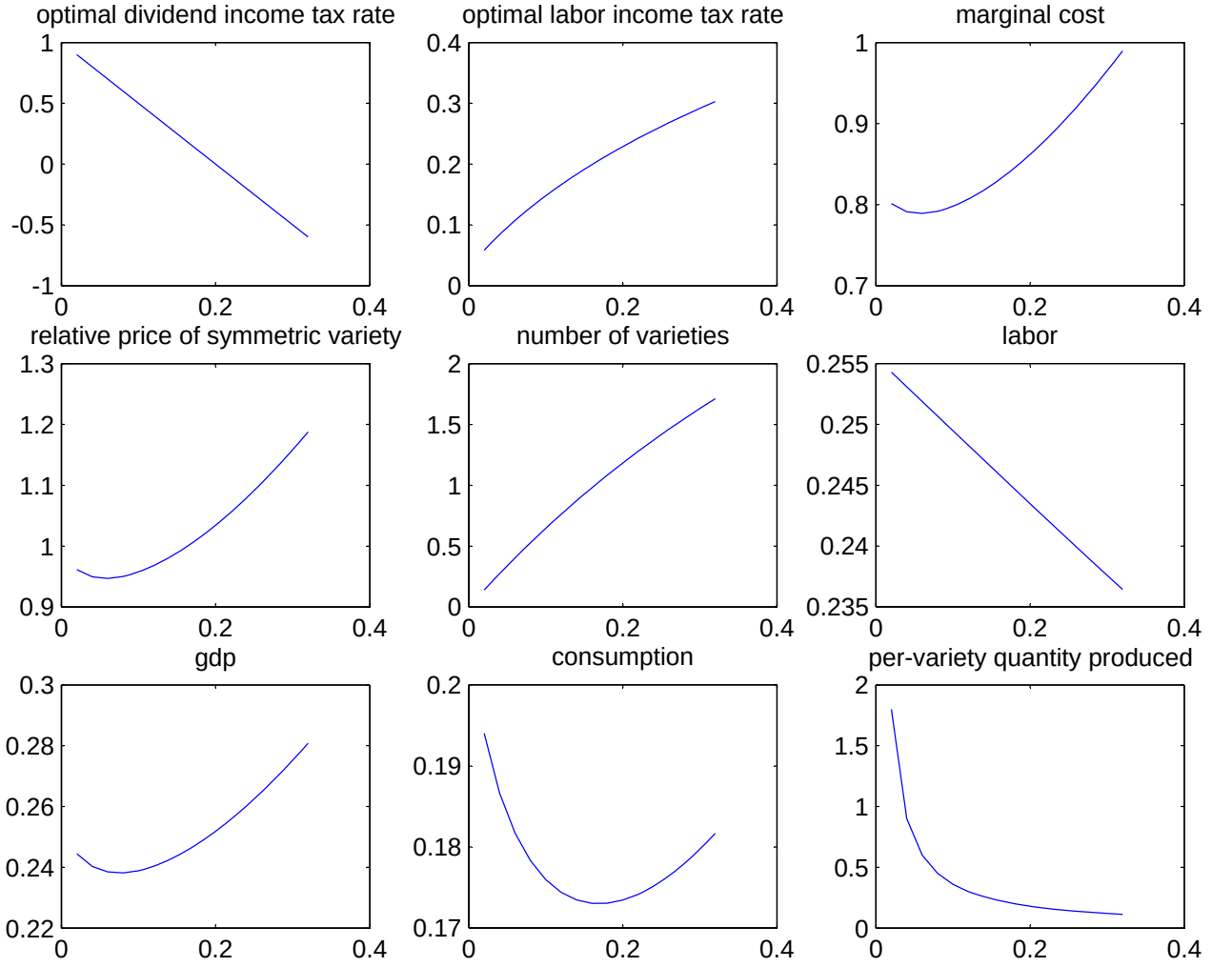


Figure 3: Long-run Ramsey equilibrium for Benassy aggregation as love of variety parameter varies, $\kappa \in [0, 0.32]$.

way because the markup is fixed in the Benassy case. Thus, profit incentives as measured by the markup are not the direct economic force shaping the long-run product creation and other equilibrium responses. Instead, the variation in τ^d across long-run Ramsey equilibria seems to be the key driver of the results, variation that does not exist in the Dixit-Stiglitz and translog aggregations. In particular, as long-run τ^d declines as the love of variety effect κ strengthens, the number of varieties increases (the middle panel of Figure 3) and total GDP increases (the lower left panel). Both of these effects occur because new product creation, N_E , is strictly increasing in κ . Although not shown in Figure 3, the rise in new product creation is quantitatively very strong, rising almost 10-fold as we vary κ between zero and 0.32. Comparing the long-run response of N_E with the long-run response of the dividend-income tax rate shown in the upper left panel, we conclude that the sharp fall in τ^d is what sparks the rise in new product creation and hence the number of varieties and GDP.

To the extent that a rise in marginal cost is akin to a decline in markups — this relation is an identity in the Dixit-Stiglitz case with a fixed number of varieties — we can interpret the other results in Figure 3 through the lens of marginal cost. Because of the fixed markup in the Benassy aggregation, a rise in the marginal cost of production directly means a rise in the relative price of a good (shown in the middle left panel). In the steady state, the real wage is by construction identical to marginal cost; we see from the middle right panel that the rise in the real wage as κ increases is associated with a decline in the Ramsey-equilibrium quantity of labor. Evidently, an inwards shift of the labor-supply function occurs as κ rises; this can be understood by recognizing that a higher κ means the love of variety effect is stronger, which, *ceteris paribus*, means households can work less (because they need to spend less) to obtain a given level of utility. GDP nonetheless rises, however, because the increase in the love of variety dominates the falls in equilibrium labor. That is, recall from the resource constraint of the economy (21) that total output measured in units of the consumption index is $\rho(N)zh$. Thus, a fall in h can indeed be associated with a rise in GDP.

In sum, the long-run Ramsey equilibrium under Benassy aggregation is different in several respects from the Ramsey equilibria under Dixit-Stiglitz aggregation and translog aggregation. In terms of short-run properties of the Ramsey equilibrium, however, the dynamics of the Benassy case are very similar to the dynamics of the Dixit-Stiglitz case.²¹ Thus, in presenting the business cycle

²¹We note a caveat here. Under Benassy aggregation, it is possible to completely shut down the variety effect by appropriately setting κ , specifically at $\kappa = 0$, in which case the relative price of a variety (and hence the variety effect) is $\rho_t = 1 \forall t$. A conjecture one may have is that any dynamic results we might find could be driven by the pure variety effect, and thus it would be of specific interest to study the dynamics of the Benassy case under $\kappa = 0$. Under the *Ramsey* equilibrium, however, we cannot consider the strict $\kappa = 0$ case because, as proven in Proposition 3, the dividend income tax in this case is 100 percent. With 100 percent dividend (profit) taxation, there is zero incentive for creation of differentiated products, and a (Ramsey) equilibrium with differentiated products does not exist; the Ramsey government prefers to have goods be perfect substitutes in this case. Thus, we cannot *completely* eliminate

properties of the Ramsey equilibrium, to which we turn next, we consider only the Dixit-Stiglitz and translog specifications.

8.3 Short-Run Optimal Policy

In Table 2, we report simulation-based moments for the Ramsey allocations and policy variables for Dixit-Stiglitz aggregation and translog aggregation. As discussed in Section 5, if both a state-contingent dividend income tax and state-contingent government debt returns are possible (recall our environment features complete government debt markets), both cannot be simultaneously pinned down along the dynamic Ramsey equilibrium. We resolve this indeterminacy by fixing the dividend income tax rate at its deterministic steady-state level and allow fluctuations in debt returns to achieve the required state-contingency of the optimal policy. By construction of the Ramsey equilibrium, all of the results presented in Table 2 are identical if we were instead to resolve the indeterminacy in the opposite way, by keeping debt returns at their steady-state level ($1/\beta$) and achieving state-contingency in the optimal policy by allowing τ^d to fluctuate. We thus report neither the dynamics of government debt returns nor the dynamics of dividend taxation because one can construct any arbitrary policy along transition paths. All other Ramsey policies and allocations, however, are determinate along transition paths.

8.3.1 Optimal Tax Rate Volatility

In terms of optimal policy dynamics, the most striking result is the very high degree of labor income tax rate variability, shown in the first row of each of the two panels of Table 2. The standard deviation of the Ramsey-optimal tax rate is greater than 2 percent, over an order of magnitude larger than in benchmark Ramsey models.²² This *lack of tax smoothing* is not a reflection of implausibly large overall volatility of the model — in both the Dixit-Stiglitz and translog cases, the coefficient of variation of GDP is about 1.6 percent, in line with empirical evidence for the U.S. economy. The extreme degree of optimal tax volatility in the face of business cycle shocks is our central result regarding Ramsey dynamics.

the variety effect in the type of Ramsey equilibrium we study. We can make it very small, however, by setting κ extremely small. For the lowest value of κ we have used, $\kappa = 0.001$, in which case relative prices N_t^κ are nearly constant over the business cycle, we still find the main tax volatility result we present below. It is true that tax volatility is about one-half what we report below for the Dixit-Stiglitz case, but it is still an order of magnitude larger than benchmark tax smoothing results in the Ramsey literature. Because we would still label this result “lack of tax smoothing,” this is the sense in which we mean that the dynamic results are no different for the Benassy case.

²²Although precise results can depend on functional form and other assumption, a benchmark value for the standard deviation of the Ramsey-optimal labor income tax rate in basic models is 0.1 percent. For more discussion, see Werning (2007), Chari and Kehoe (1999), and Arseneau and Chugh (2009).

Variable	Mean	Std. Dev.	Auto corr.	Corr(x, Y)	Corr(x, Z)	Corr(x, G)
<u>Dixit-Stiglitz aggregation</u>						
τ^h	0.2278	0.0206	0.9478	0.2786	0.9846	-0.0969
gdp	0.2519	0.0042	0.9029	1.0000	0.4132	0.8590
c	0.1734	0.0001	0.9438	-0.3084	0.5182	-0.3602
$i(=vN_E)$	0.0262	0.0018	0.8050	0.7137	0.7367	0.2747
h	0.2437	0.0051	0.9581	0.3218	-0.6944	0.6069
N	1.1839	0.0186	0.9935	0.4145	0.7004	0.3330
N_E	0.0304	0.0021	0.8000	0.7015	0.7118	0.2616
v	0.8619	0.0027	0.9935	0.4145	0.7004	0.3331
w	0.8618	0.0190	0.9411	0.4288	0.9948	0.0558
mc	0.8619	0.0027	0.9935	0.4145	0.7004	0.3331
ρ	1.0343	0.0032	0.9935	0.4145	0.7004	0.3331
μ	1.2000	0	—	—	—	—
mrs/mrt	0.6435	0.0172	0.9478	-0.2786	-0.9846	0.0969
<u>Translog aggregation</u>						
τ^h	0.1609	0.0225	0.9378	-0.0037	0.9906	-0.1055
gdp	0.2256	0.0034	0.9213	1.0000	0.1176	0.9804
c	0.1635	0.0001	0.8592	-0.2603	0.6008	-0.2252
$i(=vN_E)$	0.0125	0.0006	0.6038	0.5587	0.3945	0.3948
h	0.2496	0.0058	0.9479	0.5083	-0.7689	0.5885
N	0.6480	0.0051	0.9866	0.6684	0.5358	0.6804
N_E	0.0166	0.0008	0.5966	0.5318	0.3731	0.3666
v	0.7536	0.0016	0.9866	0.6683	0.5358	0.6803
w	0.7534	0.0157	0.9351	0.1770	0.9960	0.0764
mc	0.7536	0.0016	0.9866	0.6683	0.5358	0.6803
ρ	0.9046	0.0007	0.9866	0.6683	0.5358	0.6803
μ	1.2004	0.0016	0.9866	-0.6683	-0.5358	-0.6803
mrs/mrt	0.6989	0.0183	0.9349	0.0368	-0.9857	0.1417

Table 2: Baseline economy with endogenous product varieties. Ramsey dynamics under Dixit-Stiglitz aggregation and translog aggregation. The driving processes are shocks to both aggregate TFP and to government purchases.

We proceed in several steps and through several experiments to developing an understanding for why tax-rate volatility is optimal. First, consider the consumption-leisure equilibrium condition, (18), reproduced here for convenience:

$$-\frac{u_{ht}}{u_{ct}} = \frac{(1 - \tau_t^h)}{\mu(N_t)} z_t \rho(N_t). \quad (46)$$

Compared to the standard RBC model or to a monopolistic model with a fixed number of products, one obvious source of extra time-variation in this condition that arises solely because of product creation and destruction is N_t . As a check that it is indeed time-variation in the number of products that is driving the optimality of tax volatility, we can convert our model into one with a fixed number of products by replacing the product creation condition (19), which governs the dynamics of $N_{E,t}$, with the condition $N_{E,t} = 0 \forall t$; and by replacing the law of motion (20) with the fixed condition $N_t = \bar{N} \forall t$. We set $\bar{N} = 1$ to render this fixed-variety version of the model as close as possible to standard macro models featuring monopolistic competition. Table 3 reports the results of simulations of the Ramsey equilibrium of this alternative version of the model with the same set of realized shocks to TFP and government purchases. The first row of each panel of Table 3 shows that the standard tax-*smoothing* result indeed re-emerges — a standard deviation of τ_t^h of 0.06 percent is well within the roughly 0.1 percent volatility that is the benchmark in the literature. Thus, monopoly power in and of itself does not cause a departure from tax-smoothing. This result is consistent with those of Schmitt-Grohe and Uribe (2004a, 2004b), Siu (2004) and Chugh (2007), and confirms that it is variation in N_t that is the root cause of tax-rate variability.

Having confirmed this natural conjecture, we return to our full model and to condition (46) to try to gain further intuition for the result. In basic Ramsey models, smoothing labor tax rates over time (across states) is all about smoothing *wedges* in the consumption-leisure equilibrium condition over time (across states).²³ Adapting this idea to our environment, another conjecture is that the Ramsey equilibrium keeps time-variation in the wedge $\frac{1 - \tau_t^h}{\mu(N_t)}$ low, which we know from Corollary 1 and condition (32) is the welfare-relevant wedge; in equilibrium, this expression is equal to $\frac{MRS}{MRT} = \frac{-u_{ht}/u_{ct}}{z_t \rho(N_t)}$. The last row of each panel in Table 2 presents the business-cycle dynamics of this wedge. As is clear, this wedge also displays considerable volatility, hence tax-rate volatility seems to *not* be designed to purposefully smooth the consumption-leisure wedge over time. We think this result is surprising, especially in light of the recent results of Arseneau and Chugh (2009), who show, in the context of an environment featuring search and matching frictions in the labor market, that optimal tax rate volatility *is* designed exactly to offset other wedges (in their case, wedges due to search frictions) that impinge on the consumption-leisure margin.²⁴ Thus, the goal of

²³We again refer the interested reader to a source such as Werning (2007), Chari and Kehoe (1999), or Arseneau and Chugh (2009) for further discussion of this point.

²⁴To the extent that our model here and search and matching models are broadly similar in that they both feature

Variable	Mean	Std. Dev.	Auto corr.	Corr(x, Y)	Corr(x, Z)	Corr(x, G)
<u>Dixit-Stiglitz aggregation</u>						
τ^h	0.1380	0.0006	0.9444	0.8563	0.4640	0.8632
gdp	0.2363	0.0046	0.9322	1.0000	0.8380	0.5193
c	0.1892	0.0020	0.9259	0.7228	0.9814	-0.1668
h	0.2363	0.0025	0.9452	0.2329	-0.2944	0.9445
mrs/mrt	0.7185	0.0007	0.9288	-0.7661	-0.9921	0.1052
<u>Translog aggregation</u>						
τ^h	0.1320	0.0006	0.9445	0.8239	0.4105	0.8950
gdp	0.2335	0.0045	0.9324	1.0000	0.8387	0.5183
c	0.1864	0.0019	0.9263	0.7177	0.9797	-0.1749
h	0.2492	0.0027	0.9451	0.2361	-0.2903	0.9461
mrs/mrt	0.7685	0.0008	0.9284	-0.7589	-0.9905	0.1168

Table 3: Economy with fixed number of products, $N = 1$. Ramsey dynamics under Dixit-Stiglitz aggregation and translog aggregation. The driving processes are shocks to both aggregate TFP and to government purchases.

tax-rate volatility is *not* as simple as (nearly) completely offsetting another time-varying component of the wedge.

The next conjecture we assess is that the presence of the pre-existing *stock* of products, N_t , in the wedge condition in and of itself is the root cause of optimal tax-rate variability. That is, the wedge $\frac{1-\tau_t^h}{\mu(N_t)}$ is partially predetermined through the markup $\mu(N_t)$, which depends on the (pre-determined) number of available varieties.²⁵ Strictly speaking, this conjecture applies only to the translog aggregation, in which, as presented above, $\mu(N_t) = 1 + \frac{1}{\sigma N_t}$. In the Dixit-Stiglitz aggregation, the markup is $\mu_t = \frac{\theta}{\theta-1}$, independent of N_t . Nonetheless, it is informative to assess what happens to tax-rate variability by reducing the importance of the pre-determined stock of varieties.

To test this idea, we increase the product turnover rate δ . Increasing δ means that the flow of new products, $N_{E,t}$, constitutes a larger share of the total number of varieties at any point in time. Table 4 repeats the experiments reported in Table 2 for the case of high product turnover.

fixed costs of “entry” and hence forward-looking entry decisions, our conjecture was that the basic intuition underlying the Arseneau and Chugh (2009) result would carry over to our environment; it does not.

²⁵In contrast, in Arseneau and Chugh (2009), the search component of the wedge is not pre-determined. Rather, it is contemporaneously-determined (shadow) prices that reflect the severity of search frictions that impinge on the consumption-leisure margin.

We set $\delta = 0.90$ for the Dixit-Stiglitz case and $\delta = 0.55$ for the translog case, both well above our baseline calibration of $\delta = 0.025$.²⁶ As the first row of each panel of Table 4 shows, labor tax rate volatility is dramatically lower in both cases compared to the experiments presented in Table 2. In the Dixit-Stiglitz case, the standard deviation of the tax rate, at 0.2 percent, is not far from the Ramsey literature’s benchmark value of roughly 0.1 percent volatility. The welfare-relevant wedge in the consumption-leisure equilibrium condition also displays very low volatility, as shown in the last row of the upper panel of Table 4.

Thus, we interpret the result in the high-product-turnover environment as consistent with tax- and wedge-smoothing. This conclusion is further warranted by the fact that in this calibration of the model, overall volatility of the economy, as measured by the coefficient of variation of GDP, is nearly three times as high as in the baseline calibration. Thus, in terms of the volatilities of the tax rate and the wedge *relative to* that of total GDP, one should read the volatility results in the first row and last row of the upper panel of Table 4 as being “too high” by a factor of three.²⁷

In the translog aggregation, tax rate volatility does not decline by nearly as much as in the Dixit-Stiglitz case — its cyclical standard deviation is still 0.6 percent, as shown in the first row of the lower panel of Table 4. However, this degree of variability does support *wedge* smoothing — the standard deviation of the welfare-relevant wedge in the consumption-leisure margin is very low, at about 0.20 percent (the last row in Table 4). Just as in the Dixit-Stiglitz case with rapid product turnover, however, observe that overall volatility of the economy, as measured by the coefficient of variation of GDP, is nearly three times as high compared to the baseline calibration. We thus again interpret the results in the translog economy with high product turnover as consistent with tax- and wedge-smoothing.

It is important to note that in the limit as $\delta \rightarrow 1$, a *static* model of product turnover, similar to that studied in Jaimovich and Floetotto (2008), emerges. That is, with $\delta = 1$, product varieties have no long-lived aspect to them. Rather, with $\delta = 1$, the economy’s entire set of products is created anew every period, making the product space of the economy a static outcome.²⁸ Based on this observation and the results of our experiments, we can conclude that the crucial feature of the economy that disrupts tax-smoothing is the presence of a large *stock* in the consumption-leisure

²⁶In the translog case, we encountered numerical instability in computing dynamic decision rules for larger values of δ .

²⁷To enhance comparability with the simulation results presented earlier, we chose in the experiments underlying Table 4 to keep all parameters other than δ fixed, including those of the driving processes. We could alternatively re-calibrate the dispersions of the shocks to TFP and government spending in order to match the volatility of GDP. Because we are conducting a linear approximation, it is clear that the coefficients of variation of all variables would simply scale down by the same factor.

²⁸This idea is even clearer in the model variation we consider in Section 9, in which newly-developed products are produced and sold in markets in the same period as development, rather than with a one period delay.

Variable	Mean	Std. Dev.	Auto corr.	Corr(x, Y)	Corr(x, Z)	Corr(x, G)
<u>Dixit-Stiglitz aggregation</u>						
τ^h	0.6867	0.0020	0.9297	0.9823	0.1225	0.9879
gdp	0.0840	0.0035	0.9526	1.0000	0.2517	0.9638
c	0.0230	0.0003	0.9334	-0.1583	0.8950	-0.3957
$i(= vN_E)$	0.0116	0.0005	0.9557	0.9656	0.4759	0.8639
h	0.2474	0.0090	0.9365	0.8086	-0.3208	0.9339
N	0.0045	0.0001	0.9492	0.9205	0.4405	0.8190
N_E	0.1216	0.2041	0.9997	-0.0619	0.0892	-0.0877
v	0.2832	0.0019	0.9491	0.9202	0.4406	0.8189
w	0.2831	0.0066	0.9449	0.4655	0.9646	0.2334
mc	0.2832	0.0019	0.9491	0.9202	0.4406	0.8189
ρ	0.3398	0.0023	0.9491	0.9202	0.4406	0.8189
μ	1.2000	0	—	—	—	—
mrs/mrt	0.2613	0.0036	0.9323	-0.2701	-0.9963	-0.0271
<u>Translog aggregation</u>						
τ^h	0.7329	0.0062	0.9333	0.7682	0.7202	0.6681
gdp	0.1059	0.0047	0.9589	1.0000	0.1488	0.9868
c	0.0122	0.0001	0.7788	0.1259	0.8436	-0.0182
$i(= vN_E)$	0.0179	0.0008	0.9734	0.9820	0.2966	0.9397
h	0.2355	0.0082	0.9194	0.7575	-0.4810	0.8490
N	0.0813	0.0016	0.9616	0.9447	0.2771	0.8956
N_E	0.1035	0.0021	0.8880	0.9671	0.3041	0.9377
v	0.1732	0.0047	0.9616	0.9447	0.2771	0.8956
w	0.1732	0.0066	0.9614	0.7467	0.7099	0.6361
mc	0.1732	0.0047	0.9616	0.9447	0.2771	0.8956
ρ	0.4497	0.0070	0.9616	0.9445	0.2772	0.8955
μ	2.5985	0.0310	0.9616	-0.9443	-0.2772	-0.8953
mrs/mrt	0.1029	0.0023	0.9330	-0.3569	-0.9712	-0.2207

Table 4: High product turnover economy. Ramsey dynamics under Dixit-Stiglitz aggregation and translog aggregation with high rate of product turnover: $\delta = 0.90$ for the Dixit-Stiglitz case and $\delta = 0.55$ for the translog case. The driving processes are shocks to both aggregate TFP and to government purchases.

equilibrium condition. We know of no other model environment in which a *stock* appears in the wedge in the consumption-leisure condition. This feature of the environment cannot be the entire mechanism behind tax-rate volatility, however, because, as noted above, this feature arises only in the translog aggregation; in the Dixit-Stiglitz aggregation, the stock N_t does not enter the wedge expression because the markup is constant. Nonetheless, increasing the rate of product turnover restores the optimality of tax smoothing in the Dixit-Stiglitz case, too.

Further insight about the importance of δ for short-run tax volatility can also be gained by considering how the *long-run* elasticity of the number of new products N_E with respect to the labor income tax rate τ^h varies as δ changes. Appendix F derives the long-run elasticity of N_E with respect to τ^h , and Figure 4 plots this elasticity (in absolute value terms) as a function of the rate of product turnover δ . In both the Dixit-Stiglitz and translog aggregations, the elasticity rises very sharply as δ rises (recall that our benchmark calibration is $\delta = 0.025$). A natural conjecture is that variations in $N_{E,t}$ are crucial for cyclical stabilization of the economy (because $N_{E,t}$ directly reflects state-contingent product creation decisions). To the extent that intuition for short-run dynamics can be based on long-run equilibrium responses, the results in Figure 4 suggest that as N_E becomes more sensitive to τ^h , lower tax-rate volatility is needed in order for the Ramsey government to achieve a desired (cyclical) movement in the allocation N_E .²⁹

As $\delta \rightarrow 1$, the flow of new product creation N_E in every period approaches the flow of product destruction. Based on this observation and the results of the preceding experiments, another margin of adjustment that will be investigated in future work is how the presence of an *endogenous product destruction* decision affects optimal tax rate volatility. Endogenous product destruction is not a feature of the basic BGM framework. However, allowing such an adjustment mechanism in the private sector may be relevant for the tax volatility result based on the following logic. With a very low δ , only a small fraction of varieties are removed from the economy every period. Depending on the state of the business cycle, a social planner, and hence the Ramsey government, might prefer to remove more than the fixed fraction δ of varieties from the economy. Tax-rate volatility may thus be designed to, loosely speaking, “remove” some varieties from the economy.³⁰ For a high fixed value of δ , this possible motive for volatility in the tax rate is no longer present, and this is exactly the case for which we obtain tax smoothing in our quantitative experiments. We emphasize this is

²⁹The reason why we think that it is plausible to gain intuition for business cycle dynamics based on steady state outcomes is that business cycles in DSGE models are fairly linear phenomena around deterministic steady states, even when approximated nonlinearly.

³⁰The possibility that tax volatility is partially standing in for the lack of endogeneity of product destruction is not an instance of incompleteness of the tax system. The fact that the product destruction rate is fixed does not affect the *ability* of the government to *create* a wedge in an MRS/MRT pair that already exists in the equilibrium of the economy. Thus, while quantitative predictions may change by allowing private-market participants additional endogenous sources of adjustment, this has nothing to do with the degree of completeness of the tax system.

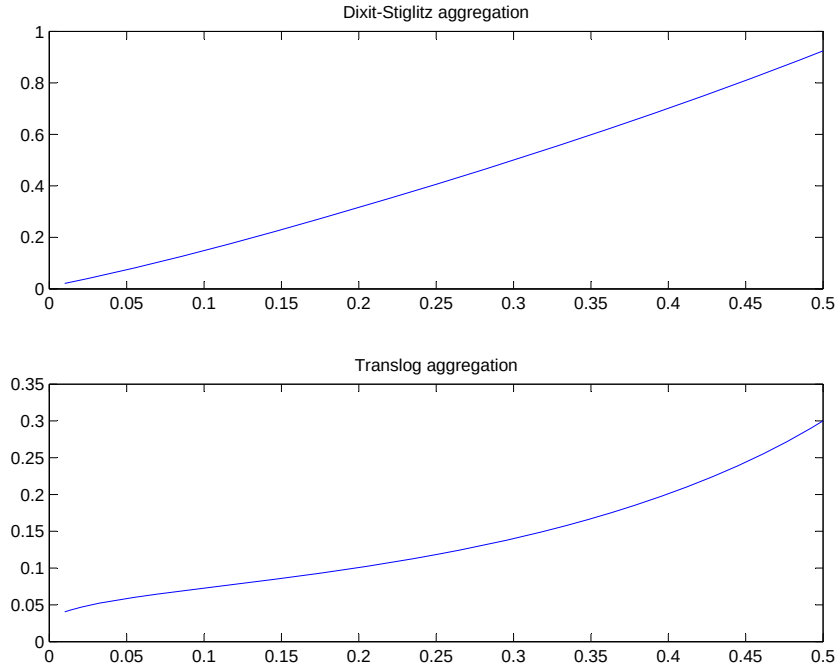


Figure 4: Long-run elasticity of N_E with respect to τ^h in Ramsey equilibria of Dixit-Stiglitz economy and translog economy, as δ varies in $\delta \in [0.01, 0.50]$.

only a conjecture at this point; based on our results, though, it is clear that tax-rate volatility is quite sensitive to the rate of product obsolescence.

8.3.2 Dynamics of Real Variables

We briefly comment on some other aspects of the Ramsey dynamics based on the baseline experiments reported in Table 2. The following observations apply to both the Dixit-Stiglitz and translog cases. First, note that gross investment, which in this economy is measured as the flow of new products created adjusted by relative prices ($v_t N_{E,t}$), is more volatile than GDP. We do not consider our work here a “data-matching” exercise, but we note that, as in BGM, this result is consistent with the idea that the stock of products is akin to the stock of physical capital in a standard RBC model. Also consistent with the dynamics predicted by a basic RBC economy is that the volatility of consumption is smaller than that of GDP — although in our case, counterfactually too low. Finally, the coefficient of variation of total hours worked, h , is roughly the same as (though slightly higher than) the coefficient of variation of GDP; in the data as reported by, say, King and Rebelo (1999), the relative volatility of these series is approximately one. Broadly, then, we view the dynamics of the Ramsey equilibrium as being in line with the data, if not a perfect match. This inspires some confidence that any insights our model delivers regarding optimal tax-rate dynamics,

which is a main focus of the study, are not obscured by wildly counterfactual predictions along other dimensions.

9 Robustness and Further Quantitative Experiments

We investigate two other model variations in order to further explore our central result of the optimality of tax-rate volatility.

9.1 Immediate Delivery to Market

Another mechanism for reducing the pre-determined aspect of the stock of product varieties is to alter the law of motion (20) to eliminate the one-period delay between product creation and availability in the goods market. We now imagine that the first time a new product created in period t is produced and sold is in period t , rather than with a one period lag as in the baseline model. Furthermore, we assume that a new variety will be produced with certainty at least once — that is, the turnover shock δ does not affect a newly-created product. The ways in which the baseline model modifies in the presence of this immediate-delivery-to-market assumption are straightforward, so we only briefly sketch the differences.

The problem of the representative firm is still to maximize (7), except now with respect to the law of motion for product varieties

$$N_t = (1 - \delta)N_{t-1} + N_{E,t}, \quad (47)$$

which captures the idea of immediate delivery to market. Optimization with respect to N_t and $N_{E,t}$ yields the production creation condition

$$mc_t f_{E,t} = (1 - \tau_t^d)(\rho_t - mc_t)q_t + (1 - \delta)E_t \left\{ \Xi_{t+1|t} mc_{t+1} f_{E,t+1} \right\}, \quad (48)$$

which should be compared with the product creation condition (12) from the baseline model.³¹ The immediate-delivery-to-market assumption does not change the firm's optimal pricing rules because pricing decisions are static.

On the household side, we keep the timing of stock demand decisions in line with the immediate delivery to market assumption by modifying the flow budget constraint from the baseline model to

$$c_t + v_t x_t N_t + b_{t+1} = (1 - \tau_t^h)w_t h_t + R_t b_t + v_t x_{t-1}(1 - \delta)N_{t-1} + (1 - \tau_t^d)d_t x_t N_t. \quad (49)$$

This budget constraint embodies the following timing of events: at the start of period t , the household owns a mutual fund of the N_{t-1} varieties that were produced in period $t - 1$. A fraction

³¹In deriving this modified product creation condition, we again impose in equilibrium that $P_{t+1} = P_t \forall t$.

δ of these varieties disappear from the market, and the market value of the shares of defunct varieties is zero. Hence, the period- t market value of the household's start-of-period shareholdings is $v_t x_{t-1} (1-\delta) N_{t-1}$. In stock-market trading of period t , the activities of ongoing and newly-created varieties are, in equilibrium, financed — that is, total stock-market purchases are $v_t x_t ((1-\delta) N_{t-1} + N_{E,t})$. Finally, following production and sales of the N_t varieties in goods markets, dividends are paid at the end of period t .

With this timing of events, household optimization with respect to x_t yields the stock demand condition

$$v_t = (1 - \tau_t^d) d_t + (1 - \delta) E_t \left\{ \Xi_{t+1|t} v_{t+1} \right\}, \quad (50)$$

which should be compared with the stock demand condition (5). Comparing (50) with the product creation condition (48) shows that, indeed, the timings of the altered stock-demand condition and the altered product creation condition remain aligned, as they were in the baseline model.

Using the same logic as presented in Appendix B, we can represent the product creation condition more generally as

$$\rho(N_t) f_{E,t} = (1 - \tau_t^d) (\mu(N_t) - 1) \left(\frac{c_t + g_t}{N_t} \right) + (1 - \delta) E_t \left\{ \frac{\mu(N_t)}{\mu(N_{t+1})} \rho(N_{t+1}) f_{E,t+1} \right\}. \quad (51)$$

In the definition of the private-sector equilibrium, this condition obviously replaces the product creation (19), and the law of motion (47) replaces (20). As we show in Appendix D, the PVIC modifies only slightly — specifically, the time-zero assets of the household now do *not* include a dividend income term because dividend income is no longer earned off a *predetermined* stock of products. Apart from these modifications, the definition of the private-sector equilibrium is as in Section 2.5.

Table 5 presents simulation results for this version of the model. The parameter values used for the simulations reported here are the same as those underlying the experiments in Table 2; in particular, we have reset the rate of obsolescence to its baseline value, $\delta = 0.025$. Comparing the results in Table 5 to those in Table 2, we see that the unconditional means of policy and allocation variables are virtually identical, as would be expected. In terms of the central focus of our study, the immediate-delivery-to-market timing implies somewhat lower optimal tax volatility: at about one percent in both the Dixit-Stiglitz and translog cases, the standard deviation of the tax rate is roughly half that in the baseline economy. The last rows of each panel show that fluctuations in the Ramsey-optimal wedge are also roughly two-thirds as large as in the baseline model. In terms of overall volatility, the standard deviation of GDP, at about two percent, is a bit higher than the baseline model but still in line with empirical evidence.

Our conclusion from this experiment is that, although shortening the lag between the introduction of a new good and actual production and sales of that good reduces the degree of optimal

Variable	Mean	Std. Dev.	Auto corr.	Corr(x, Y)	Corr(x, Z)	Corr(x, G)
<u>Dixit-Stiglitz aggregation</u>						
τ^h	0.2265	0.0097	0.9162	0.6694	0.9441	-0.2887
gdp	0.2540	0.0071	0.9037	1.0000	0.8524	0.4701
c	0.1750	0.0014	0.9913	0.6556	0.8653	0.0233
$i(=vN_E)$	0.0266	0.0039	0.8291	0.8229	0.8132	0.0533
h	0.2440	0.0039	0.9204	0.4480	-0.0311	0.7918
N	1.2257	0.0419	0.9939	0.5654	0.7577	0.0747
N_E	0.0306	0.0044	0.8246	0.8048	0.7861	0.0502
v	0.8678	0.0059	0.9939	0.5654	0.7576	0.0748
w	0.8678	0.0219	0.9548	0.8180	0.9850	0.0273
mc	0.8678	0.0059	0.9939	0.5654	0.7576	0.0748
ρ	1.0414	0.0071	0.9939	0.5654	0.7576	0.0748
μ	1.2000	0	—	—	—	—
mrs/mrt	0.6450	0.0123	0.9561	-0.6884	-0.9606	0.1811
<u>Translog aggregation</u>						
τ^h	0.1610	0.0086	0.9109	0.6709	0.9613	-0.2384
gdp	0.2244	0.0053	0.8876	1.0000	0.8237	0.5084
c	0.1615	0.0015	0.9822	0.6665	0.9455	-0.0810
$i(=vN_E)$	0.0130	0.0021	0.7069	0.6683	0.6451	-0.0291
h	0.2494	0.0035	0.9257	0.4017	-0.1471	0.8695
N	0.6996	0.0195	0.9880	0.5792	0.8515	-0.0244
N_E	0.0175	0.0028	0.7010	0.6446	0.6086	-0.0282
v	0.7437	0.0058	0.9879	0.5790	0.8512	-0.0243
w	0.7437	0.0198	0.9591	0.7782	0.9885	-0.0005
mc	0.7437	0.0058	0.9879	0.5790	0.8512	-0.0243
ρ	0.9001	0.0026	0.9879	0.5789	0.8511	-0.0242
μ	1.2104	0.0059	0.9879	-0.5789	-0.8510	0.0242
mrs/mrt	0.6936	0.0101	0.9388	-0.6681	-0.9693	0.2113

Table 5: Immediate delivery to market economy. Ramsey dynamics under Dixit-Stiglitz aggregation and translog aggregation. The driving processes are shocks to both aggregate TFP and to government purchases.

tax-rate volatility, the crucial feature for restoring tax-smoothing is still a more rapid rate of product turnover, as the earlier experiments suggested. Indeed, even in this immediate-delivery-to-market environment, increasing δ re-establishes tax-smoothing and wedge-smoothing as the optimal policy; for brevity, we do not report these results. Condition (48) shows that in the limiting case of complete product turnover every period ($\delta = 1$), a purely static condition governs the number of product varieties in any period, an idea also captured by the model of Jaimovich and Floetotto (2008).

9.2 Product Creation Subsidies

Finally, we allow the Ramsey government access to another tax instrument to test whether its availability may re-instate labor income tax smoothing. Returning to our baseline model, we now suppose the government can subsidize the cost of product creation at the proportional rate τ_t^s . BGM2 also considered this subsidy, and, in the context of their deterministic analysis, it was a powerful instrument. Our hypothesis here is that this tax instrument will *not* re-instate the optimality of tax smoothing for reasons shown by BGM2 themselves. Specifically, they showed that under either Dixit-Stiglitz aggregation or Benassy aggregation, such a subsidy would be *constant* across time and states. It is only with translog aggregation, in which markups fluctuate, that BGM2 showed that a product creation subsidy would be expected to vary over time. Because the lack of labor tax smoothing in our analysis arises for *all three* types of aggregation and because we do not see how, in the Dixit-Stiglitz and Benassy cases at least, a *constant* product creation subsidy over time could affect the *volatility* of labor tax rates we have already found, our conjecture is that time-varying creation subsidies are not what tax-rate volatility is proxying for.

Nevertheless, to test the idea, suppose that firms receive a subsidy on the cost of developing a new variety. Inclusive of the subsidy, the cost to the firm of developing a new variety is thus $(1 - \tau_t^s)mc_t f_{E,t}$. It is easy to show that the product creation condition now includes the subsidy as part of the cost of and asset value of a variety. The product creation condition is now

$$(1 - \tau_t^s)\rho(N_t)f_{E,t} = (1 - \delta)E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d)\mu(N_t) \left(1 - \frac{1}{\mu(N_{t+1})} \right) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + (1 - \tau_{t+1}^s) \frac{\mu(N_t)}{\mu(N_{t+1})} \rho(N_{t+1})f_{E,t+1} \right] \right\}, \quad (52)$$

which is straightforward to derive after introducing $(1 - \tau_t^s)$ in the firm profit function (7). For simplicity, we assume that after the government transfers these subsidies to firms, they are subsequently rebated lump-sum by firms to households. Aggregate product subsidies $\tau_t^s mc_t f_{E,t} N_{E,t}$ are thus an income item for households in (2) and an expenditure item for the government in (16). No other equilibrium conditions are altered.³² In line with our hybrid approach to the Ramsey

³²When BGM2 consider product creation subsidies, they modify the resource constraint to account for the subsidy. We think it more natural to include these subsidies in the government's and households' budget constraints — that

Variable	Mean	Std. Dev.	Auto corr.	Corr(x, Y)	Corr(x, Z)	Corr(x, G)
<u>Translog aggregation</u>						
τ^h	0.1779	0.0116	0.5135	-0.8797	-0.5225	0.8247
gdp	0.2342	0.0184	0.6291	1.0000	0.8429	-0.5222
c	0.1576	0.0005	0.7091	-0.9312	-0.7504	0.4931
h	0.2591	0.0183	0.6254	0.9977	0.8065	-0.5755
τ^s	-1.0000	0.0025	0.8672	-0.0313	-0.0755	0.0425
mrs/mrt	0.6832	0.0102	0.5875	0.9105	0.5553	-0.8223

Table 6: Economy with product creation subsidy. Ramsey dynamics under translog aggregation. The driving processes are shocks to both aggregate TFP and to government purchases.

problem, in computing the Ramsey equilibrium, we simply compute a first-order condition directly with respect to τ_t^s ; note that τ_t^s affects the dynamic Ramsey program through the period- t budget constraint as well as the period- $(t - 1)$ (forward-looking) product creation condition.

In the Ramsey steady state, the creation subsidy is redundant with respect to the dividend income tax. That is, in the margin (33), both the dividend income tax and the creation subsidy appear. We resolve the indeterminacy by shutting down the dividend income tax entirely. Thus, the product creation subsidy corrects both long-run and short-run distortions in product creation. As derived in BGM2, the long-run subsidy is characterized by

$$1 - \tau^s = \frac{\mu(N) - 1}{\epsilon(N)}, \quad (53)$$

and the short-run subsidy is characterized by

$$1 - \tau_t^s = \frac{\mu(N_{t+1})}{\mu(N_t)}(1 - \tau_{t-1}^s). \quad (54)$$

Together, these two expressions characterize both the level and the time-path of the subsidy; these results carry over to the full Ramsey environment.

Table 6 presents dynamics, for translog aggregation, of key variables in the presence of a possibly time-varying creation subsidy. As proven by BGM2, the long-run level of the subsidy is $\tau^s = -1$ with translog aggregation due to the configuration of monopoly incentives and the love of variety effect. The fluctuations in the subsidy rate over the business cycle turn out to be miniscule, as the next to last row of Table 6 shows.

For our purposes, the main result is that the labor income tax rate, shown in the first row of Table 6, is still quite volatile. Even though its standard deviation is halved relative to the case is, in equilibrium, they are transfers between the government and the private sector — rather than include them in a description of the technology of the economy.

without creation subsidies (refer back to Table 2), it is still over an order of magnitude more volatile than benchmark Ramsey results. We do not even report results in Table 6 for the Dixit-Stiglitz or Benassy aggregations because, as we noted above, the creation subsidy is constant in these cases and labor tax rate dynamics continue to be the same as in the baseline model. Indeed, for the Dixit-Stiglitz aggregation, the results are *identical* to those reported in Table 2 because the product creation subsidy is always zero, both in the long run and the short run.

The two extensions considered in this section thus go in the direction of reinstating tax smoothing as the optimal policy. However, neither restores the result as typically understood in the Ramsey literature.

10 Conclusion

In this paper, we studied optimal fiscal policy in an environment in which product varieties are the result of purposeful, forward-looking investment decisions by firms. One main result is that the long-run optimal dividend income tax rate is positive in the most empirically relevant and intuitively appealing version of the model. Depending on the form of variety aggregation, it is also possible that a long-run dividend income *subsidy* is instead optimal; however, the optimality of a strictly *zero* dividend tax is non-generic. Our second main result is that keeping labor income tax rates virtually constant in the face of business cycle shocks, a prescription of all simple Ramsey models, is *not* optimal. The key feature of the economy driving this result is the inherently *dynamic* nature of product creation decisions. If product creation decisions were instead purely static phenomena, tax smoothing is optimal. A broad message this latter result suggests is that, while modeling a fluctuating number of products may be important for insights regarding some business-cycle questions, modeling a fluctuating number of products driven by *dynamic* profit incentives may deliver yet further insights.

As we have discussed, one important extension to our work is to allow for an endogenous product destruction margin. In addition to independent interest in developing the BGM framework along these lines for the analysis of business cycles, this margin could be important for the dynamics of optimal policy prescriptions. Another interesting research question that would build on our work here is to introduce money demand motives in order to study jointly optimal fiscal and monetary policy, in the tradition begun by Lucas and Stokey (1983) and Chari, Christiano, and Kehoe (1991). A complementary line of research would be to consider the business-cycle as well as optimal-policy implications of product turnover dynamics in conjunction with the employment turnover dynamics articulated by modern search and matching frameworks; regarding pure business cycle dynamics, Shao and Silos (2009) is an early work in this direction.

The most promising next step we plan to take is studying optimal fiscal policy in a model

along these latter lines: one featuring product turnover dynamics in conjunction with employment turnover dynamics. Bringing these two ideas together is natural given our results here and other recent results in the Ramsey literature. We showed here that tax smoothing is not optimal in the presence of long-lived product stocks. Arseneau and Chugh (2009) showed that tax smoothing is not optimal in the most common specification of business-cycle search frameworks. A natural next question, given that it is useful to think of both products and employment as long-lived phenomena, is the nature of optimal policy in such an environment.

References

- AIYAGARI, S. RAO, ALBERT MARCET, THOMAS J. SARGENT, AND JUHA SEPPALA. 2002. "Optimal Taxation without State-Contingent Debt." *Journal of Political Economy*, Vol. 110, pp. 1220-1254.
- ARMENTER, ROC. 2008. "A Note on Incomplete Factor Taxation." *Journal of Public Economics*, Vol. 92, pp. 2275-2281.
- ARSENEAU, DAVID M. AND SANJAY K. CHUGH. 2008. "Optimal Fiscal and Monetary Policy with Costly Wage Bargaining." *Journal of Monetary Economics*, Vol. 55, pp. 1401-1414.
- ARSENEAU, DAVID M. AND SANJAY K. CHUGH. 2009. "Tax Smoothing in Frictional Labor Markets." International Finance Discussion Paper no. 965, Board of Governors of the Federal Reserve System.
- ARSENEAU, DAVID M., SANJAY K. CHUGH, AND ANDRE KURMANN. 2008. "Optimal Capital Taxation in an Economy with Capital Allocation Frictions." University of Maryland.
- BENASSY, JEAN-PASCAL. 1996. "Taste for Variety and Optimum Production Patterns in Monopolistic Competition." *Economics Letters*, Vol. 52, pp. 41-47.
- BENASSY, JEAN-PASCAL. 1998. "Is There Always Too Little Research in Endogenous Growth With Expanding Product Variety?." *European Economic Review*, Vol. 42, pp. 61-69.
- BILBIE, FLORIN, FABIO GHIRONI, AND MARC J. MELITZ. 2007. "Endogenous Entry, Product Variety, and Business Cycles." NBER Working Paper 13646.
- BILBIE, FLORIN, FABIO GHIRONI, AND MARC J. MELITZ. 2008. "Monopoly Power and Endogenous Product Variety: Distortions and Remedies." NBER Working Paper 14383.
- BRODA, CHRISTIAN AND DAVID E. WEINSTEIN. 2007. "Product Creation and Destruction: Evidence and Price Implications." NBER Working Paper 13041.
- CHAMLEY, CHRISTOPHE. 1986. "The Welfare Cost of Capital Income Taxation in a Growing Economy." *Econometrica*, Vol. 54, pp. 607-622.
- CHARI, V.V., LAWRENCE CHRISTIANO, AND PATRICK KEHOE. 1991. "Optimal Fiscal and Monetary Policy: Some Recent Results." *Journal of Money, Credit, and Banking*, Vol. 23, pp. 519-539.
- CHARI V. V., AND PATRICK J. KEHOE. 1999. "Optimal Fiscal and Monetary Policy. In *Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford, Vol. 1C. Elsevier.
- CHUGH, SANJAY K. 2007. "Optimal Inflation Persistence: Ramsey Taxation with Capital and Habits." *Journal of Monetary Economics*, Vol. 54, pp. 1809-1836.
- DIXIT, AVINASH K. AND JOSEPH E. STIGLITZ. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review*, Vol. 67, pp. 297-308.

- FAIA, ESTER. 2008. "Optimal Monetary Policy Rules with Labor Market Frictions." *Journal of Economic Dynamics and Control*, Vol. 32, pp. 1357-1370.
- FEENSTRA, ROBERT C. 2003. "A Homothetic Utility Function for Monopolistic Competition Models, Without Constant Price Elasticity." *Economics Letters*, Vol. 78, pp. 79-86.
- GROSSMAN, GENE M. AND ELHANAN HELPMAN. 1991. *Innovation and Growth in the Global Economy*. MIT Press.
- JAIMOVICH, NIR AND MAX FLOETOTTO. 2008. "Firm Dynamics, Markup Variations, and the Business Cycle." *Journal of Monetary Economics*, Vol. 55, pp. 1208-1252.
- JUDD, KENNETH L. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, Vol. 28, pp. 59-83.
- KHAN, AUBHIK, ROBERT G. KING AND ALEXANDER L. WOLMAN. 2003. "Optimal Monetary Policy." *Review of Economic Studies*, Vol. 70, pp. 825-860.
- KING, ROBERT G. AND SERGIO T. REBELO. 1999. "Resuscitating Real Business Cycles. In *Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford, Vol. 1B. Elsevier.
- LUCAS, ROBERT E. JR. AND NANCY L. STOKEY. 1983. "Optimal Fiscal and Monetary Policy in an Economy Without Capital." *Journal of Monetary Economics*, Vol. 12, pp. 55-93.
- MAS-COLELL, ANDREW, MICHAEL D. WHINSTON, AND JERRY R. GREEN. 1995. *Microeconomic Theory*. Oxford University Press.
- ROMER, PAUL M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, Vol. 98, pp. S71-102.
- SCHMITT-GROHE, STEPHANIE AND MARTIN URIBE. 2004a. "Optimal Fiscal and Monetary Policy Under Imperfect Competition." *Journal of Macroeconomics*, Vol. 26, pp. 183-209.
- SCHMITT-GROHE, STEPHANIE AND MARTIN URIBE. 2004b. "Optimal Fiscal and Monetary Policy Under Sticky Prices." *Journal of Economic Theory*, Vol. 114, pp. 198-230.
- SCHMITT-GROHE, STEPHANIE AND MARTIN URIBE. 2004c. "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function." *Journal of Economic Dynamics and Control*, Vol. 28, pp. 755-775.
- SHAO, ENCHUAN AND PEDRO SILOS. 2009. "Firm Entry and Labor Market Dynamics." Federal Reserve Bank of Atlanta.
- SIU, HENRY E. 2004. "Optimal Fiscal and Monetary Policy with Sticky Prices." *Journal of Monetary Economics*, Vol. 51, pp. 576-607.
- THOMAS, CARLOS. 2007. "Search and Matching Frictions and Optimal Monetary Policy ." London School of Economics.

A Derivation of Pricing Equation

The nominal lifetime profit function is

$$\max_{\tilde{P}_t, N_{t+1}, N_{E,t}} E_0 \sum_{t=0}^{\infty} \Xi_{t|0} \left[(1 - \tau_t^d)(\tilde{P}_t - P_t mc_t) N_t q(\tilde{P}_t) - P_t mc_t f_{E,t} N_{E,t} \right] \quad (55)$$

where $\tilde{P}_t$ is the nominal price of a symmetric variety and $q(\tilde{P}_t)$ is the associated demand function for that symmetric variety. In period t , the firm faces the law of motion for the number of varieties it produces and sells,

$$N_{t+1} = (1 - \delta)(N_t + N_{E,t}). \quad (56)$$

The first order condition with respect to the nominal price $\tilde{P}_t$ of a symmetric variety is

$$N_t q(\tilde{P}_t) + \tilde{P}_t N_t q'(\tilde{P}_t) - P_t mc_t N_t q'(\tilde{P}_t) = 0. \quad (57)$$

From this, we now proceed to derive a simple representation of the optimal pricing condition (for a symmetric variety). Canceling the N_t terms and rearranging gives

$$\tilde{P}_t q'(\tilde{P}_t) = P_t mc_t q'(\tilde{P}_t) - q(\tilde{P}_t). \quad (58)$$

Rearranging further,

$$\tilde{P}_t = P_t mc_t - \frac{1}{\frac{1}{q(\tilde{P}_t)} q'(\tilde{P}_t)}. \quad (59)$$

Dividing by P_t ,

$$\frac{\tilde{P}_t}{P_t} = mc_t - \frac{1}{\frac{P_t}{q(\tilde{P}_t)} q'(\tilde{P}_t)}. \quad (60)$$

Multiplying and dividing the denominator of the second term on the right hand side by $\tilde{P}_t$,

$$\frac{\tilde{P}_t}{P_t} = mc_t - \frac{1}{\frac{\tilde{P}_t}{q(\tilde{P}_t)} q'(\tilde{P}_t) \frac{P_t}{\tilde{P}_t}}. \quad (61)$$

Rewriting,

$$\frac{\tilde{P}_t}{P_t} = mc_t - \frac{\tilde{P}_t}{P_t} \frac{1}{\frac{\tilde{P}_t}{q(\tilde{P}_t)} q'(\tilde{P}_t)}. \quad (62)$$

Defining $\zeta_t \equiv \frac{\tilde{P}_t}{q(\tilde{P}_t)} q'(\tilde{P}_t)$ as the price elasticity of demand for a symmetric variety, we have

$$\frac{\tilde{P}_t}{P_t} \left(1 + \frac{1}{\zeta_t} \right) = mc_t. \quad (63)$$

The optimal relative price of a symmetric variety is thus

$$\frac{\tilde{P}_t}{P_t} = \left(\frac{\zeta_t}{1 + \zeta_t} \right) mc_t, \quad (64)$$

which is an endogenously-time-varying markup over real marginal cost. Denoting by ρ_t the optimal relative price of a symmetric variety, $\rho_t \equiv \tilde{P}_t/P_t$, and by μ_t the gross markup, $\mu_t \equiv \frac{\zeta_t}{1+\zeta_t}$,

$$\rho_t = \mu_t mc_t. \tag{65}$$

B Private-Sector Equilibrium

The most straightforward definition of equilibrium is that it is a set of 15 endogenous equilibrium processes $\{c_t, h_t, h_t^E, h_t^G, N_{t+1}, N_{E,t}, w_t, v_t, \mu_t, \rho_t, \xi_t, mc_t, q_t, d_t, b_t\}_{t=0}^\infty$, for given exogenous processes $\{z_t, g_t, \tau_t^h, \tau_t^d, f_{E,t}\}$, that satisfy the conditions listed below. Recall that h_t^G is the *per-product* level of employment in the goods-producing sector, hence $h_t^G N_t$ is the total quantity of labor employed in the goods-producing sector; and h_t^E is the per-product level of employment in the product-development sector, hence $h_t^E N_{E,t}$ is the total quantity of labor employed in the product-development sector. In the following, the one-period-ahead stochastic discount factor is given by $\Xi_{t+1|t} = \frac{\beta u_c(c_{t+1}, h_{t+1})}{u_c(c_t, h_t)}$.

There are fifteen sequences of conditions describe the private-sector equilibrium; most of these have already been derived in the description of the model, but we collect them all here for convenience.

Consumption-leisure optimality condition:

$$-\frac{u_h(c_t, h_t)}{u_c(c_t, h_t)} = (1 - \tau_t^h)w_t \quad (66)$$

The relation between the marginal cost of production and the real wage:

$$mc_t = \frac{w_t}{z_t} \quad (67)$$

Stock-demand condition:

$$v_t = (1 - \delta)E_t \left\{ \Xi_{t+1|t} \left[v_{t+1} + (1 - \tau_{t+1}^d) d_{t+1} \right] \right\} \quad (68)$$

Optimal pricing of a symmetric variety:

$$\rho_t = \mu_t mc_t \quad (69)$$

Relation between gross markup and price elasticity of demand:

$$\mu_t = \frac{\xi_t}{1 + \xi_t} \quad (70)$$

Product creation condition:

$$mc_t f_{E,t} = (1 - \delta)E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d) (\rho_{t+1} - mc_{t+1}) q_{t+1} + mc_{t+1} f_{E,t+1} \right] \right\} \quad (71)$$

Law of motion for the number of product varieties:

$$N_{t+1} = (1 - \delta)(N_t + N_{E,t}) \quad (72)$$

Per-variety quantity of goods produced:

$$q_t = \frac{c_t + g_t}{N_t \rho_t} \quad (73)$$

Aggregate resource constraint:

$$c_t + g_t + \rho_t N_{E,t} f_{E,t} = \rho_t z_t h_t \quad (74)$$

Condition that pins down hours worked in “product-development sector:”

$$h_t^E = \frac{f_{E,t}}{z_t} \quad (75)$$

Aggregate hours:

$$h_t = h_t^E N_{E,t} + h_t^G N_t \quad (76)$$

Per-variety dividends:

$$d_t = (\rho_t - mc_t) q_t \quad (77)$$

Government budget constraint:

$$\tau_t^h w_t h_t + \tau_t^d d_t + b_t = g_t + R_t b_{t-1} \quad (78)$$

Finally, we have not yet described the parametric forms we adopt for preferences, which directly imply particular equilibrium expressions for the gross markup and the variety effect. As in BGM, which provides details, we explore three different functional forms: for the Dixit-Stiglitz CES case, the goods-market markup is given by

$$\mu_t = \frac{\theta}{\theta - 1}, \quad (79)$$

and the love-of-variety effect is described by

$$\rho_t = N_t^{\mu_t - 1} \left(= N_t^{\frac{1}{\theta - 1}} \right). \quad (80)$$

For the Feenstra (2003) translog expenditure function, the goods-market markup is given by

$$\mu_t = 1 + \frac{1}{\sigma N_t}, \quad (81)$$

and the love-of-variety effect is described by

$$\rho_t = e^{-\frac{1}{2} \frac{\tilde{N} - N_t}{\sigma \tilde{N} N_t}}. \quad (82)$$

And for the CES Benassy (1996) specification, which disentangles the elasticity of substitution between varieties from the love of variety effect, the goods-market markup is given by

$$\mu_t = \frac{\theta}{\theta - 1}, \quad (83)$$

and the love-of-variety effect is described by

$$\rho_t = N_t^\kappa. \quad (84)$$

Clearly, if $\kappa = \frac{1}{\theta - 1}$, the Benassy specification is identical to the Dixit-Stiglitz specification,

B.1 Compact Representation of Equilibrium

To characterize the equilibrium more compactly, as presented in Section 2.5, we combine the above conditions in the following way. First, emphasize the general dependence of the relative price on the number of varieties as $\rho(N_t)$; this function of course depends on which functional form for preferences we are considering. Similarly, emphasize the (possible) dependence of the goods-market markup on the number of varieties as $\mu(N_t)$.

As we showed in Appendix ??, $v_t = mc_t f_{E,t}$; this implies that the stock-demand condition and the product creation condition are identical in equilibrium. So we drop the stock-demand condition and proceed to carry around only the product creation condition, in which obviously the definition of per-firm dividend payments, $d_t = (\rho(N_t) - mc_t)q_t$, already appears.

Next, substitute $q_t = \frac{c_t + g_t}{N_t \rho(N_t)}$ in the product creation condition to express it as

$$\frac{\rho(N_t)}{\mu(N_t)} f_{E,t} = (1-\delta) E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d) \left(\rho(N_{t+1}) - \frac{\rho(N_{t+1})}{\mu(N_{t+1})} \right) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1} \rho(N_{t+1})} \right) + \frac{\rho(N_{t+1})}{\mu(N_{t+1})} f_{E,t+1} \right] \right\}, \quad (85)$$

in which we have also made the substitution $mc_t = \frac{\rho(N_t)}{\mu(N_t)}$. Canceling $\rho(N_{t+1})$ terms, we have

$$\frac{\rho(N_t)}{\mu(N_t)} f_{E,t} = (1-\delta) E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d) \left(1 - \frac{1}{\mu(N_{t+1})} \right) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \frac{\rho(N_{t+1})}{\mu(N_{t+1})} f_{E,t+1} \right] \right\}. \quad (86)$$

Multiplying by $\mu(N_t)$, we have a compact representation of the product creation condition

$$\rho(N_t) f_{E,t} = (1-\delta) E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d) \mu(N_t) \left(1 - \frac{1}{\mu(N_{t+1})} \right) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \frac{\mu(N_t)}{\mu(N_{t+1})} \rho(N_{t+1}) f_{E,t+1} \right] \right\}, \quad (87)$$

which is condition (19) in the text.

To obtain the model's static optimality condition, use the relation $w_t = z_t mc_t$ in the consumption-leisure optimality condition, and then use the relation $mc_t = \frac{\rho(N_t)}{\mu(N_t)}$ to eliminate marginal cost. The resulting expression for the equilibrium consumption-leisure characterization is

$$-\frac{u_{ht}}{u_{ct}} = \frac{(1 - \tau_t^h)}{\mu(N_t)} z_t \rho(N_t), \quad (88)$$

which is condition (18) in the text.

B.2 Aggregate Resource Constraint

To derive the representation of the aggregate resource constraint above and presented in (21), sum the flow household budget constraint and the flow government budget constraint, which gives

$$c_t + g_t + v_t N_{E,t} = w_t h_t + d_t N_t. \quad (89)$$

Substitute into this expression the equilibrium expression for (per-product) dividends, $d_t = (\rho_t - mc_t)q_t$,

$$c_t + g_t + v_t N_{E,t} = w_t h_t + (\rho_t - mc_t)q_t N_t. \quad (90)$$

Next, apply the definition of the quantity of each variety, $q_t = \frac{c_t + g_t}{N_t \rho_t}$; canceling terms leaves

$$c_t + g_t + v_t N_{E,t} = w_t h_t + (\rho_t - mc_t) \left(\frac{c_t + g_t}{\rho_t} \right). \quad (91)$$

Next, using the condition $v_t = mc_t f_{E,t}$,

$$c_t + g_t + mc_t f_{E,t} N_{E,t} = w_t h_t + (\rho_t - mc_t) \left(\frac{c_t + g_t}{\rho_t} \right); \quad (92)$$

and expressing marginal cost as $mc_t = \rho_t / \mu_t$,

$$c_t + g_t + \frac{\rho_t}{\mu_t} f_{E,t} N_{E,t} = w_t h_t + \left(\rho_t - \frac{\rho_t}{\mu_t} \right) \left(\frac{c_t + g_t}{\rho_t} \right). \quad (93)$$

Canceling terms on the right hand-side,

$$c_t + g_t + \frac{\rho_t}{\mu_t} f_{E,t} N_{E,t} = w_t h_t + \left(1 - \frac{1}{\mu_t} \right) (c_t + g_t), \quad (94)$$

and canceling the $(c_t + g_t)$ that appears on both sides,

$$\frac{1}{\mu_t} (c_t + g_t) + \frac{\rho_t}{\mu_t} N_{E,t} f_{E,t} = w_t h_t. \quad (95)$$

Next, recognize that $w_t = z_t mc_t = z_t \frac{\rho_t}{\mu_t}$, we have

$$\frac{1}{\mu_t} (c_t + g_t) + \frac{\rho_t}{\mu_t} N_{E,t} f_{E,t} = z_t \frac{\rho_t}{\mu_t} h_t. \quad (96)$$

Finally, multiplying by μ_t gives

$$c_t + g_t + \rho_t N_{E,t} f_{E,t} = \rho_t z_t h_t, \quad (97)$$

the form of the resource constraint we use. This emphasizes the idea that ρ_t should be thought of as a primitive feature of the economy.

C Efficient Allocations

The social planning problem is to choose state-contingent functions for $\{c_t, h_t, N_{t+1}, N_{E,t}\}$ to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, h_t) \quad (98)$$

subject to

$$c_t + g_t + \rho(N_t)N_{E,t}f_{E,t} = \rho(N_t)z_t h_t \quad (99)$$

and

$$N_{t+1} = (1 - \delta)(N_t + N_{E,t}). \quad (100)$$

The social planner internalizes the effect of the number of varieties on the relative price, which we emphasize by writing $\rho(N_t)$ in the resource constraint.

Let ϕ_t denote the Lagrange multiplier on the goods-market constraint and μ_t denote the Lagrange multiplier on the law of motion for the number of product varieties. The first-order conditions with respect to c_t , h_t , $N_{E,t}$, and N_{t+1} are, respectively,

$$u_{ct} - \phi_t = 0, \quad (101)$$

$$u_{ht} + \phi_t \rho(N_t) z_t = 0, \quad (102)$$

$$-\phi_t \rho(N_t) f_{E,t} + \mu_t (1 - \delta) = 0, \quad (103)$$

and

$$-\mu_t + \beta E_t \{ \phi_{t+1} \rho'(N_{t+1}) [z_{t+1} h_{t+1} - N_{E,t+1} f_{E,t+1}] + (1 - \delta) \mu_{t+1} \} = 0. \quad (104)$$

Conditions (101) and (102) imply

$$-\frac{u_{ht}}{u_{ct}} = z_t \rho(N_t). \quad (105)$$

This is the efficiency condition (24) that appears in the main text. Note that this implies the welfare-relevant marginal rate of transformation between leisure and consumption is $z_t \rho(N_t)$, which takes into account not only the marginal product of the physical production technology, z_t , but also the love-of-variety effect, which is summarized by $\rho(N_t)$.

Solving condition (103) for μ_t and substituting $\phi_t = u_{ct}$ from condition (101), we have

$$\mu_t = \frac{u_{ct} \rho(N_t) f_{E,t}}{1 - \delta}. \quad (106)$$

Using the time- t and time- $t + 1$ versions of this expression in condition (104), we have

$$u_{ct} \rho(N_t) f_{E,t} = \beta (1 - \delta) E_t \{ u_{ct+1} [\rho'(N_{t+1}) (z_{t+1} h_{t+1} - N_{E,t+1} f_{E,t+1}) + \rho(N_{t+1}) f_{E,t+1}] \}. \quad (107)$$

We now apply several definitions and identities to simplify this expression. By the identity $f_{E,t} = z_t h_t^E$, this can be re-written as

$$u_{ct}\rho(N_t)f_{E,t} = \beta(1 - \delta)E_t \left\{ u_{ct+1} \left[\rho'(N_{t+1})(z_{t+1}h_{t+1} - N_{E,t+1}z_{t+1}h_{t+1}^E) + \rho(N_{t+1})f_{E,t+1} \right] \right\}. \quad (108)$$

Next, by the definition $h_t = h_t^G N_t + h_t^E N_{E,t}$, this can be re-written as

$$u_{ct}\rho(N_t)f_{E,t} = \beta(1 - \delta)E_t \left\{ u_{ct+1} \left[\rho'(N_{t+1})z_{t+1}N_{t+1}h_{t+1}^G + \rho(N_{t+1})f_{E,t+1} \right] \right\}, \quad (109)$$

where, recall, h_t^G is the labor hired *per firm* in the goods-producing sector. Next, use the goods-producing firm's production technology, $q_t = z_t h_t^G$, to express this as

$$u_{ct}\rho(N_t)f_{E,t} = \beta(1 - \delta)E_t \left\{ u_{ct+1} \left[\rho'(N_{t+1})q_{t+1}N_{t+1} + \rho(N_{t+1})f_{E,t+1} \right] \right\}. \quad (110)$$

By definition, $q_t = \frac{c_t + g_t}{N_t \rho(N_t)}$; using this, we can rewrite the preceding as

$$u_{ct}\rho(N_t)f_{E,t} = \beta(1 - \delta)E_t \left\{ u_{ct+1} \left[\rho'(N_{t+1}) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}\rho(N_{t+1})} \right) N_{t+1} + \rho(N_{t+1})f_{E,t+1} \right] \right\}. \quad (111)$$

The love-of-variety effect expressed in elasticity form is $\epsilon(N_t) \equiv \rho'(N_t) \frac{N_t}{\rho(N_t)}$; using this, we can again re-express the preceding as

$$u_{ct}\rho(N_t)f_{E,t} = \beta(1 - \delta)E_t \left\{ u_{ct+1} \left[\epsilon(N_{t+1}) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \rho(N_{t+1})f_{E,t+1} \right] \right\}. \quad (112)$$

Finally, using the definition of the one-period-ahead stochastic discount factor, $\Xi_{t+1|t} \equiv \frac{\beta u_{ct+1}}{u_{ct}}$, we have

$$\rho(N_t)f_{E,t} = (1 - \delta)E_t \left\{ \Xi_{t+1|t} \left[\epsilon(N_{t+1}) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \rho(N_{t+1})f_{E,t+1} \right] \right\}. \quad (113)$$

This is the dynamic efficiency condition (25) that appears in the main text. In the deterministic steady state, we have that the efficient level of product creation is characterized by

$$\frac{1}{\beta} = (1 - \delta) \left[\frac{\epsilon(N) \left(\frac{c+g}{N} \right) + \rho(N)f_E}{\rho(N)f_E} \right], \quad (114)$$

which is the long-run efficiency condition (28) that appears in the main text.

D Derivation of Present-Value Implementability Constraint

Here we derive the present-value implementability constraint (PVIC), which is a present-value version of the household's budget constraint with all prices and policies substituted out using equilibrium conditions of the model. The PVIC is expressed in terms of only allocations, a typical approach in Ramsey problems, and constrains the set of allocations that can be implemented by the Ramsey government. We first present the derivation for the baseline model, in which there is a one-period delay between product creation and sales; we then present the derivation for the model with the assumption of immediate delivery to market.

D.1 Baseline Model

Start with the household flow budget constraint,

$$c_t + v_t x_{t+1}(N_t + N_{E,t}) + b_{t+1} = (1 - \tau_t^h)w_t h_t + R_t b_t + [v_t + (1 - \tau_t^d)d_t]x_t N_t. \quad (115)$$

Multiply each term by $\beta^t u_{ct}$ (which, in equilibrium, is the shadow value to the household of a unit of wealth) and, conditional on the information set at time zero, sum the sequence of budget constraints from $t = 0 \dots \infty$ to arrive at

$$\begin{aligned} & E_0 \sum_{t=0}^{\infty} \beta^t u_{ct} c_t + E_0 \sum_{t=0}^{\infty} \beta^t u_{ct} v_t x_{t+1}(N_t + N_{E,t}) + E_0 \sum_{t=0}^{\infty} \beta^t u_{ct} b_{t+1} \\ &= E_0 \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^h) w_t h_t + E_0 \sum_{t=0}^{\infty} \beta^t u_{ct} R_t b_t + E_0 \sum_{t=0}^{\infty} \beta^t u_{ct} [v_t + (1 - \tau_t^d) d_t] x_t N_t. \end{aligned}$$

We now begin to impose equilibrium conditions on this present-value budget constraint. For ease of notation, we drop the E_0 term, but it is understood that all terms are conditional on the information set of time zero. First impose the sequence of stock-market clearing conditions $x_s = 1 \forall s$, which gives

$$\begin{aligned} & \sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ct} v_t (N_t + N_{E,t}) + \sum_{t=0}^{\infty} \beta^t u_{ct} b_{t+1} \\ &= \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^h) w_t h_t + \sum_{t=0}^{\infty} \beta^t u_{ct} R_t b_t + \sum_{t=0}^{\infty} \beta^t u_{ct} [v_t + (1 - \tau_t^d) d_t] N_t. \end{aligned}$$

Next, in the third summation on the left-hand-side, we substitute the sequence of bond Euler equations, $u_{cs} = \beta E_s \{u_{cs+1} R_{s+1}\}$, $\forall s$,

$$\begin{aligned} & \sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ct} v_t (N_t + N_{E,t}) + \sum_{t=0}^{\infty} \beta^{t+1} u_{ct+1} R_{t+1} b_{t+1} \\ &= \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^h) w_t h_t + \sum_{t=0}^{\infty} \beta^t u_{ct} R_t b_t + \sum_{t=0}^{\infty} \beta^t u_{ct} [v_t + (1 - \tau_t^d) d_t] N_t. \end{aligned}$$

Note that, by the law of iterated expectations, the conditional expectation E_s drops back to the time-zero expectation E_0 . Canceling terms in the third summation on the left-hand-side with their counterpart terms in the second summation on the right-hand-side leaves only the time-zero bond-return term,

$$\sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ct} v_t (N_t + N_{E,t}) = \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^h) w_t h_t + \sum_{t=0}^{\infty} \beta^t u_{ct} [v_t + (1 - \tau_t^d) d_t] N_t + u_{c0} R_0 b_0. \quad (116)$$

Next, in the first summation on the right-hand-side, use the sequence of consumption-leisure optimality conditions, $-u_{hs} = u_{cs}(1 - \tau_s^h)w_s$, $\forall s$, and move this summation to the left-hand-side,

$$\sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ht} h_t + \sum_{t=0}^{\infty} \beta^t u_{ct} v_t (N_t + N_{E,t}) = \sum_{t=0}^{\infty} \beta^t u_{ct} [v_t + (1 - \tau_t^d) d_t] N_t + u_{c0} R_0 b_0. \quad (117)$$

We next use the sequence of stock-pricing equations, $v_s = (1 - \delta)E_s \left\{ \frac{\beta u_{cs+1}}{u_{cs}} [v_{s+1} + (1 - \tau_{s+1}^d) d_{s+1}] \right\}$, $\forall s$, to substitute out the $u_{cs} v_s$ terms in the third summation on the left-hand-side, which yields

$$\begin{aligned} \sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ht} h_t + (1 - \delta) \sum_{t=0}^{\infty} \beta^{t+1} u_{ct+1} [v_{t+1} + (1 - \tau_{t+1}^d) d_{t+1}] (N_t + N_{E,t}) \\ = \sum_{t=0}^{\infty} \beta^t u_{ct} [v_t + (1 - \tau_t^d) d_t] N_t + u_{c0} R_0 b_0. \end{aligned}$$

Using the sequence of equilibrium laws of motion $\frac{N_{s+1}}{1 - \delta} = N_s + N_{E,s}$, $\forall s$, in the third summation on the left-hand-side, we have

$$\sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ht} h_t + \sum_{t=0}^{\infty} \beta^{t+1} u_{ct+1} [v_{t+1} + (1 - \tau_{t+1}^d) d_{t+1}] N_{t+1} = \sum_{t=0}^{\infty} \beta^t u_{ct} [v_t + (1 - \tau_t^d) d_t] N_t + u_{c0} R_0 b_0. \quad (118)$$

Canceling terms in the third summation on the left-hand-side with their counterpart terms in the summation on the right-hand-side leaves only the time-zero stock-payoff term,

$$\sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ht} h_t = u_{c0} [v_0 + (1 - \tau_0^d) d_0] N_0 + u_{c0} R_0 b_0. \quad (119)$$

Re-introducing the expectation E_0 and recognizing that in equilibrium $v_0 = mc_0 f_{E,0}$ and $d_0 = (\rho_0 - mc_0)q_0$, we arrive at the PVIC

$$\sum_{t=0}^{\infty} \beta^t [u_{ct} c_t + u_{ht} h_t] = u_{c0} [v_0 + (1 - \tau_0^d) d_0] N_0 + u_{c0} R_0 b_0. \quad (120)$$

The PVIC thus encapsulates the (present-value) household budget constraint, the sequence of stock-pricing conditions, the sequence of laws of motion of the number of product varieties, the sequence of bond-pricing conditions, and the sequence of consumption-leisure optimality conditions. Thus, the remaining equilibrium conditions — i.e., those not captured by the PVIC — are the sequence of aggregate goods resource constraints, the sequence of free-entry conditions, and the sequence of optimal pricing conditions.

D.2 Immediate Delivery to Market

Under the alternative timing, we start with the household flow budget constraint

$$c_t + v_t x_t N_t + b_{t+1} = (1 - \tau_t^h) w_t h_t + R_t b_t + v_t x_{t-1} (1 - \delta) N_{t-1} + (1 - \tau_t^d) d_t x_t N_t. \quad (121)$$

Once again multiplying by $\beta^t u_{ct}$ and, conditional on the information set at time zero, summing the sequence of budget constraints from $t = 0 \dots \infty$ (and dropping the E_0 operator because it is understood), we have

$$\begin{aligned} & \sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ct} v_t x_t N_t + \sum_{t=0}^{\infty} \beta^t u_{ct} b_{t+1} \\ &= \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^h) w_t h_t + \sum_{t=0}^{\infty} \beta^t u_{ct} R_t b_t + \sum_{t=0}^{\infty} \beta^t u_{ct} v_t x_{t-1} (1 - \delta) N_{t-1} + \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^d) x_t d_t N_t. \end{aligned}$$

Proceeding exactly as above by first imposing the sequence of stock-market clearing conditions, $x_s = 1 \forall s$; substituting the sequence of bond Euler equations; applying the law of iterated expectations; substituting the sequence of consumption-leisure optimality conditions, we arrive at

$$\sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ht} h_t + \sum_{t=0}^{\infty} \beta^t u_{ct} v_t N_t = \sum_{t=0}^{\infty} \beta^t u_{ct} v_t (1 - \delta) N_{t-1} + \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^d) d_t N_t + u_{c0} R_0 b_0. \quad (122)$$

We next use the sequence of stock-pricing equations, $v_s = (1 - \tau_s^d) d_s + (1 - \delta) E_s \left\{ \frac{\beta u_{cs+1}}{u_{cs}} v_{s+1} \right\}$, $\forall s$, to substitute in the third summation on the left hand side, which yields

$$\begin{aligned} & \sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ht} h_t + \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^d) d_t N_t + \sum_{t=0}^{\infty} \beta^{t+1} u_{ct+1} v_{t+1} (1 - \delta) N_t \\ &= \sum_{t=0}^{\infty} \beta^t u_{ct} v_t (1 - \delta) N_{t-1} + \sum_{t=0}^{\infty} \beta^t u_{ct} (1 - \tau_t^d) d_t N_t + u_{c0} R_0 b_0. \end{aligned}$$

The third summation on the left hand side exactly cancels with the second summation on the right hand side, which leaves

$$\sum_{t=0}^{\infty} \beta^t u_{ct} c_t + \sum_{t=0}^{\infty} \beta^t u_{ht} h_t + \sum_{t=0}^{\infty} \beta^{t+1} u_{ct+1} v_{t+1} (1 - \delta) N_t = \sum_{t=0}^{\infty} \beta^t u_{ct} v_t (1 - \delta) N_{t-1} + u_{c0} R_0 b_0. \quad (123)$$

Finally, the third summation on the left hand side cancels with the summation on the right hand side, leaving only the time-zero term; reintroducing the E_0 operator, we have the PVIC for the immediate-delivery-to-market model,

$$E_0 \sum_{t=0}^{\infty} \beta^t [u_{ct} c_t + u_{ht} h_t] = u_{c0} v_0 (1 - \delta) N_{-1} + u_{c0} R_0 b_0. \quad (124)$$

E Optimal Long-Run Dividend Tax

Here we prove Proposition 3. As stated in Section 5, the Ramsey problem is to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, h_t) \quad (125)$$

subject to the sequence of technological constraints

$$c_t + g_t + \rho(N_t)N_{E,t}f_{E,t} = \rho(N_t)z_t h_t \quad (126)$$

and

$$N_{t+1} = (1 - \delta)(N_t + N_{E,t}), \quad (127)$$

the sequence of product creation conditions

$$\rho(N_t)f_{E,t} = (1 - \delta)E_t \left\{ \Xi_{t+1|t} \left[(1 - \tau_{t+1}^d)\mu(N_t) \left(1 - \frac{1}{\mu(N_{t+1})} \right) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \frac{\mu(N_t)}{\mu(N_{t+1})} \rho(N_{t+1})f_{E,t+1} \right] \right\}, \quad (128)$$

and the PVIC

$$E_0 \sum_{t=0}^{\infty} \beta^t [u_{ct}c_t + u_{ht}h_t] = u_{c0} [R_b b_0 + (v_0 + (1 - \tau_0^d)d_0)N_0]. \quad (129)$$

The Ramsey choice variables are c_t , h_t , N_{t+1} , $N_{E,t}$, and τ_t^d for $t > 1$. Associate the sequences of multipliers $\lambda_{1,t}$, $\lambda_{2,t}$, $\lambda_{3,t}$ with the first three sequences of constraints, and the multiplier ξ with the PVIC. Although we of course must consider the fully dynamic Ramsey problem to consider any aspect of the Ramsey equilibrium, our analytical results are for only the deterministic Ramsey steady state. Thus, here we can suppose the environment is deterministic and drop all expectations operators.

The first-order condition with respect to τ_t^d , which, note, requires differentiating through the time- $t - 1$ production creation condition (recall that we focus on only timeless Ramsey allocations — which is akin to ignoring the time $t = 0$ Ramsey first-order condition, as we do), immediately implies that $\lambda_3 = 0$ in the deterministic Ramsey steady state. This is a very useful result because it greatly simplifies the analysis of the rest of the Ramsey steady state. Intuitively, what the result $\lambda_3 = 0$ says is that in the Ramsey equilibrium (though, note, not in any arbitrary equilibrium), the product creation condition does not constrain the allocation. Stated another way, the Ramsey government arranges allocations so that in the long-run the product creation margin is at its *efficient* setting. In what follows we rely on the long-run result that $\lambda_3 = 0$.

To prove results for the long-run optimal dividend income tax, we need to consider only the Ramsey first-order conditions with respect to N_{t+1} and $N_{E,t}$. These first-order conditions are, respectively,

$$-\lambda_{2,t} + \beta [\lambda_{1,t+1}\rho'(N_{t+1})(z_{t+1}h_{t+1} - N_{E,t+1}f_{E,t+1}) + (1 - \delta)\lambda_{2,t+1}] = 0 \quad (130)$$

and

$$-\lambda_{1,t}\rho(N_t)f_{E,t} + (1 - \delta)\lambda_{2,t} = 0. \quad (131)$$

We have ignored any derivatives through the product creation condition because, as just derived above, the Lagrange multiplier on this constraint is zero in the deterministic steady state. Then, using exactly the same set of algebraic manipulations as in Appendix C, these two conditions can be expressed as

$$\lambda_{1,t}\rho(N_t)f_{E,t} = \beta(1 - \delta) \left\{ \lambda_{1,t+1} \left[\epsilon(N_{t+1}) \left(\frac{c_{t+1} + g_{t+1}}{N_{t+1}} \right) + \rho(N_{t+1})f_{E,t+1} \right] \right\}. \quad (132)$$

In the deterministic steady state, we have the Ramsey-optimal level of product creation is characterized by

$$\frac{1}{\beta} = (1 - \delta) \left[\frac{\epsilon(N) \left(\frac{c+g}{N} \right) + \rho(N)f_E}{\rho(N)f_E} \right], \quad (133)$$

which is exactly the long-run efficiency condition (28) that appears in the main text. Thus, the Ramsey equilibrium achieves the Pareto optimum along the product-creation margin in the long-run.

To decentralize this, refer to the deterministic steady-state version of the product creation condition:

$$\frac{1}{\beta} = (1 - \delta) \left[\frac{(1 - \tau^d)(\mu(N) - 1) \left(\frac{c+g}{N} \right) + \rho(N)f_E}{\rho(N)f_E} \right]. \quad (134)$$

Comparing these last two expressions, it is clear that in the long-run Ramsey equilibrium, the dividend income tax rate is characterized by

$$1 - \tau^d = \frac{\epsilon(N)}{\mu(N) - 1}, \quad (135)$$

or

$$\tau^d = 1 - \frac{\epsilon(N)}{\mu(N) - 1}. \quad (136)$$

This result is exactly as in BGM2, even though they do not consider a full Ramsey-optimal taxation problem.

The Ramsey first-order conditions with respect to c_t and h_t are, respectively,

$$u_{ct} - \lambda_{1,t} + \xi [u_{cct}c_t + u_{ct}] = 0 \quad (137)$$

and

$$u_{ht} + \lambda_{1,t}\rho(N_t)z_t + \xi [u_{hht}h_t + u_{ht}] = 0. \quad (138)$$

These conditions, except for the presence of the variety effect, are the same as in a baseline Ramsey model — for example, see Chari and Kehoe (1999). Combining these, we have

$$\frac{-u_{ht}/u_{ct}}{z_t\rho(N_t)} = 1 + \xi \left\{ 1 + \frac{u_{cct}c_t}{u_{ct}} + \frac{1}{z_t\rho(N_t)} \left[\frac{u_{hht}h_t}{u_{ct}} + \frac{u_{ht}}{u_{ct}} \right] \right\}, \quad (139)$$

which is expression (34) in the main text.

F Long-Run Elasticity of New Products with Respect to Taxes

To derive the long-run elasticity of N_E with respect to τ^h , start with the deterministic steady state versions of the private-sector equilibrium conditions, but ignoring the government budget constraint, presented in Section 2.5:

$$-\frac{u_h}{u_c} = (1 - \tau^h) \frac{\rho(N)}{\mu(N)}, \quad (140)$$

$$\rho(N) = \beta(1 - \delta) \left[(1 - \tau^d)(\mu(N) - 1) \left(\frac{c}{N} \right) + \rho(N) \right], \quad (141)$$

$$\delta N = (1 - \delta)N_E, \quad (142)$$

and

$$c + \rho(N)N_E = \rho(N)h. \quad (143)$$

In this presentation of the equilibrium, we assume zero government spending without loss of generality. Also, recall the labor market clearing condition $h = N_E h^E + N h^G$, and we have normalized $h^E = 1$.

After tedious but straightforward substitutions, we can represent the long-run equilibrium as the single conditions

$$F(N_E, \tau^h; :) \equiv \frac{1}{\beta(1 - \delta)} - 1 + (1 - \tau^d)h^G \left[1 + (1 - \tau^h)\rho \left(\frac{1 - \delta}{\delta} N_E \right) \frac{u_c(c(N_E))}{u_h(h(N_E))} \right] = 0, \quad (144)$$

in which, following from goods-market and labor-market clearing, we have $c(N_E) = \rho \left(\frac{1 - \delta}{\delta} N_E \right)$ and $h(N_E) = N_E + \frac{1 - \delta}{\delta} N_E h^G$. Using the Implicit Function Theorem and again through tedious but straightforward algebra, we have

$$\begin{aligned} \frac{\partial N_E}{\partial \tau^h} &= - \frac{F_{\tau^h}}{F_{N_E}} \\ &= \frac{\rho(N)u_c(c)}{(1 - \tau^h) \left[\frac{(1 - \delta)\rho'(N)u_c(c)}{\delta} + \frac{(1 - \delta)\rho(N)u_{cc}(c)(\rho'(N)N h^G + \rho(N))}{\delta} - \frac{\rho(N)u_c(c)u_{hh}(h) \left(1 + \frac{(1 - \delta)h^G}{\delta} \right)}{u_h(h)} \right]}. \end{aligned}$$

The elasticity is then of course $\epsilon_{N_E, \tau^h} = \frac{\partial N_E}{\partial \tau^h} \frac{\tau^h}{N_E}$.