Essays on Taxation and Real Estate

Essays on Taxation and Real Estate

A dissertation presented

by

Nathaniel E. Hipsman

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

Harvard University

Cambridge, Massachusetts

April 2018

Dissertation Advisor:
Professor Stefanie Stantcheva

Author:
Nathaniel E. Hipsman

Essays on Taxation and Real Estate

Abstract

This dissertation consists of three independent chapters. In the first chapter, "Optimal Taxation in Overlapping Generations Economies with Aggregate Risk," I ask how governments should leverage available policy instruments to raise revenue and share aggregate risk across generations. I analytically derive two new, opposing considerations in this setting, in addition to the classic desire to smooth distortions, and then consider numerical applications to three policy problems: financing of wars, intergenerational sharing of productivity risk, and intergenerational redistribution of trend productivity growth. The second chapter, "Race and Home Price Appreciation in the United States: 1992–2012," investigates the the extent to which home price appreciation is related to homeowner race. I link transaction-level data on home sale prices to mortgage application-level data from the Home Mortgage Disclosure Act on applicant demographics to answer this question. I find that black (Hispanic) homeowners experience appreciation that is 1 to 2.5 (0.5 to 1.5) percentage points lower than their white counterparts. In the third chapter, "The Effect of State Income Taxes on Home Values: Evidence from a Border Pair Study," I estimate the elasticity of home prices with respect to the net-of-state-income-tax rate using several variants of a border pair strategy. I find suggestive evidence that this elasticity is positive and large, but also discuss several concerns with drawing too strong a conclusion.

Contents

List of Tables

List of Figures

Acknowledgments

I wish to thank many individuals without whom this dissertation would not have been possible. First and foremost, I am grateful to my advisors—Emmanuel Farhi, Edward Glaeser, David Laibson, N. Gregory Mankiw, and Stefanie Stantcheva—for their continuing assistance and guidance over several years; each played a unique and pivotal role in helping me create the dissertation presented today. Second, I am very much indebted to my parents, Linda and Steven Hipsman, for their invaluable emotional and financial support throughout my time in graduate school and all preceding education. Third, I acknowledge many other Harvard professors and seminar participants, as well as several anonymous journal referees, for helpful comments during preliminary presentations of this work. Finally, I thank my classmates Ashvin Gandhi and Shoshana Vasserman for many detailed conversations that led to substantial improvements in the work contained in this dissertation.

Introduction

This dissertation consists of three independent chapters. The first, "Optimal Taxation in Overlapping Generations Economies with Aggregate Risk," asks how governments should leverage available policy instruments to raise revenue and share aggregate risk across generations. I address this question by developing a framework for analyzing optimal taxation in economies with overlapping generations (OLG) and stochastic government spending and productivity. I derive two new, opposing considerations in addition to the classic desire to smooth distortions. First, such economies lack Ricardian equivalence. This encourages governments to run balanced budgets, since deficits drive up interest rates and therefore future tax distortions. Second, the social planner has a redistributive motive across generations and thus faces an equity/efficiency tradeoff. I consider applications to three policy problems: financing of wars, intergenerational sharing of productivity risk, and intergenerational redistribution of trend productivity growth. I find that optimal policy in the first application features partial tax smoothing, with substantially higher labor taxes when government spending is high, but also substantial autocovariance of the labor tax rate. I demonstrate in the latter two applications the optimality of a Social Security program with procyclical benefits; this program is larger if trend productivity growth is more rapid and the planner is more inequality-averse.

The second, "Race and Home Price Appreciation in the United States: 1992–2012," asks whether homeowners of different racial and ethnic groups experience different rates of home price appreciation. I study this question by linking transaction-level repeat sale home price data to demographic information disclosed under the Home Mortgage Disclosure Act. I find that, after flexibly accounting for differences in income, commuting zone, and the timing of purchase and sale, black homeowners experience appreciation that is 2.5 percentage points lower than their white counterparts; the analogous Hispanic-white gap is 1.5 percentage points. I then document that homeowners of different races and ethnicities, but the same incomes, select homes in neighborhoods with very different characteristics along dimensions that are correlated with home price appreciation, such as racial and income makeup. However, I show that flexibly controlling for all of these factors still leaves an unexplained black-white appreciation gap of 1 percentage point, and a Hispanic-white gap of about 0.5 percentage points. Finally, I show that these gaps vary widely across commuting zones, but in ways that are generally not correlated with other salient economic attributes of the commuting zones, such as interracial income mobility gaps.

Finally, the third, "The Effect of State Income Taxes on Home Values: Evidence from a Border Pair Study," estimates the elasticity of home prices with respect to the net-of-state-income-tax rate using several variants of a border pair strategy. I find suggestive evidence that this elasticity is positive and large. However, the estimates are highly imprecise, and event studies produce mixed results. Calibrating a calculation of the marginal value of public funds for state income taxes to the elasticities suggested by this study shows that ignoring such general equilibrium effects of taxes can lead to large errors in the calculation, and therefore more work should be done to obtain more precise estimates of the elasticity.

Chapters 1 and 3 share an interest in developing better tax policies; Chapters 2 and 3 share an interest in home pricing. More broadly however, all three chapters strongly contribute to the formulation of optimal policy, whether directly through theoretical or empirical analysis of taxation, or indirectly through concern for interracial equity regarding an important economic asset.

Chapter 1

Optimal Taxation in Overlapping Generations Economies with Aggregate Risk

1.1 Introduction

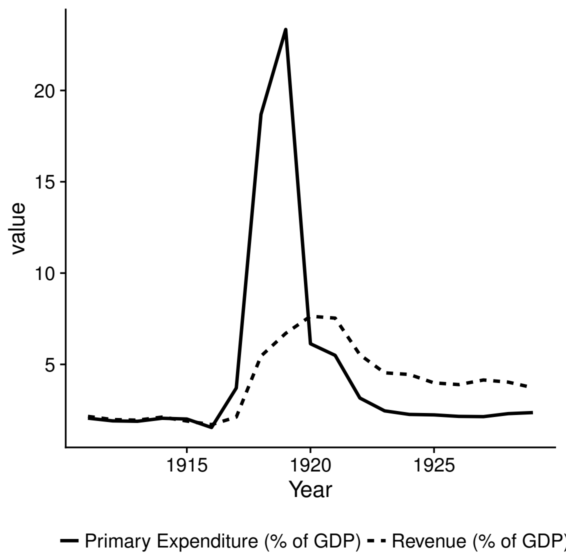
Most economic policies have differential effects on different generations. On the revenue side of the ledger, labor taxation's incidence falls primarily on the young, who make up a large portion of the labor force, while capital taxation's falls primarily on the old, who possess most of the economy's wealth. On the spending side, Medicare and Social Security explicitly redistribute from younger to older citizens, while education subsidies, child tax credits, and the Earned Income Tax Credit primarily benefit younger generations, even if such intergenerational redistribution is not those programs' intended purpose. Important national debates are currently being argued over possible reforms to many of these programs, most visibly Social Security, which faces a funding crisis in the near future. These debates point up a broad economic policy question: How should governments leverage available policy instruments—labor taxes, capital taxes, lump-sum transfers, and debt—to raise revenue and share aggregate risk across generations? This paper develops a framework for analyzing many variants of this problem.

Optimal distortionary taxation in a stochastic, general equilibrium model is not a novel question but instead a classic macroeconomic policy topic, having been extensively analyzed for economies featuring infinitely-lived households. In such models, intergenerational risk sharing is not a concern, leading to exclusive focus on efficiently raising revenue to fund exogenous, stochastic government spending, usually conceptualized as mandatory wars. This problem is often described as the "Ramsey taxation problem." The major contributions to this literature, detailed below, all derive different versions of the same policy guidance: Labor taxation should be nearly constant over time, or "smooth." Specifically, if complete insurance markets are available, then government budget shocks should be perfectly insured, leaving marginal distortionary costs of taxation constant over time; if not, then governments should borrow the full value of budget shocks, leaving marginal distortionary costs of labor taxation to follow a risk-adjusted random walk.

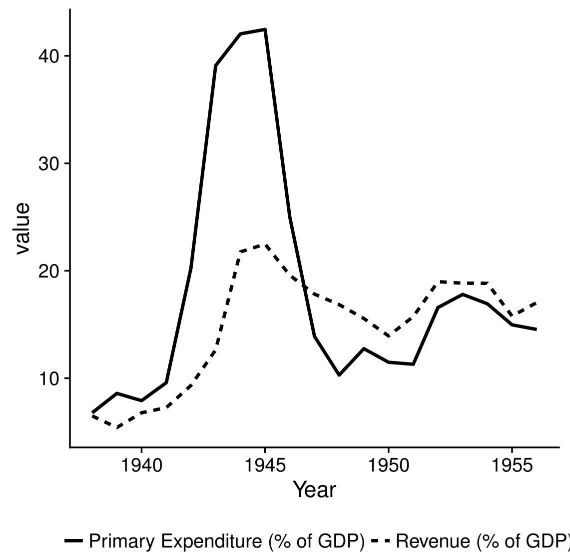
This result predicts well the empirical behavior of governments in response to large government spending shocks. Figure 1.1 shows the fiscal response to the largest spending shocks on record—the two World Wars—in the only two countries for which good data exists through the period. While the sample size is small and the data quite noisy, the graphs suggest that developed countries at least approximately follow the random walk advice of an incomplete markets Ramsey model. Indeed, one cannot reject a unit root for the path of revenue as a percentage of GDP.[1]

However, the real world is not, in fact, populated by one infinitely-lived cohort of

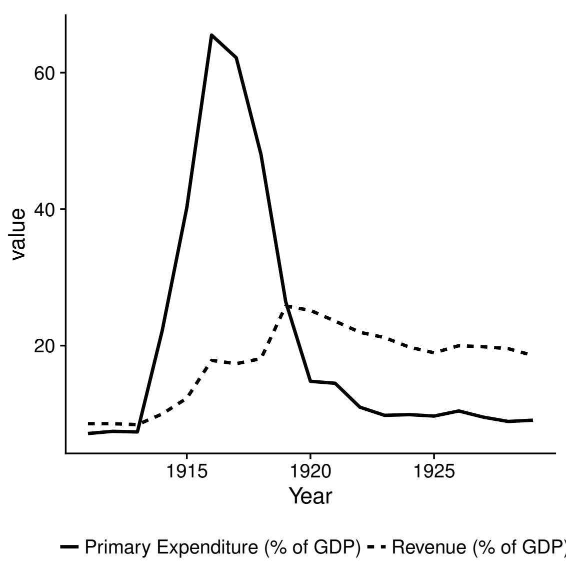
[1] We will see in the numerical section that for certain values of the parameters one can obtain a similar path for optimal policy in an OLG model.



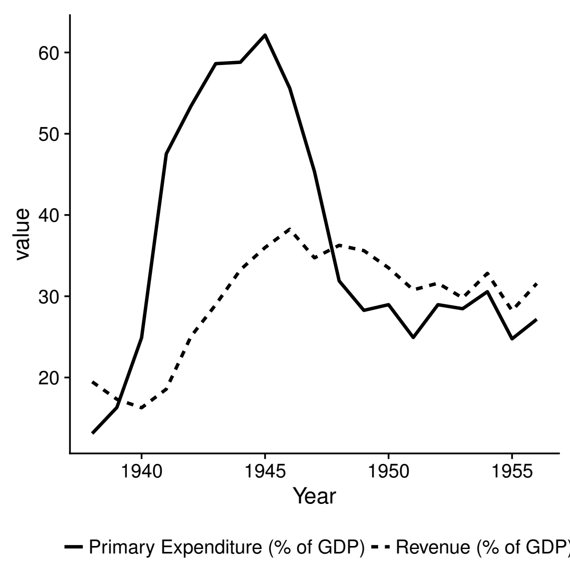
(a) World War I, U.S.



(b) World War II, U.S.



(c) World War I, U.K.



(d) World War II, U.K.

Figure 1.1: Primary expenditure and revenue as a percentage of GDP for four large government spending shocks. Data source: Mauro et al. (2015)

households but rather by a series of overlapping generations, and these recommended policies are highly inequitable across those generations. Insurance against government spending shocks places that risk disproportionately on older generations, who are likely to have accumulated the most assets and therefore be the most likely counterparties to the insurance contracts. Government borrowing instead places that risk disproportionately on younger and unborn generations, who will face steeper taxes in the future.

Furthermore, substantial portions of developed countries' government budgets are spent not on purchases of final goods and services, but on old-age transfers. The U.S., for example, spent 41% of its primary[2] federal budget on Social Security and Medicare in 2016 (Center on Budget and Policy Priorities, 2017). The optimality of such spending, and the extensive tax revenue and debt required to fund it, cannot be analyzed within the context of a representative-agent model. More broadly, the Ramsey taxation literature can safely ignore economies without government purchases but with other sorts of aggregate risk, so long as there is a representative agent. Since the welfare theorems apply and there is no heterogeneity, the *laissez-faire* competitive equilibrium is the only Pareto optimal outcome. However, when the economy consists of a series of overlapping generations, there is a continuum of Pareto optimal outcomes, some of which may be preferable to the *laissez-faire* competitive equilibrium from a distribution perspective.

The framework developed herein addresses these limitations of the existing optimal dynamic taxation literature by making a single, important change to the model: replacing the homogeneous, infinite-horizon households with a series of overlapping generations (OLG). These OLG households have a parameterized bequest motive and (weakly negative) lower limit on their net worth at end of life. Since such an economy no longer

[2]excluding interest on the debt

features a representative agent, one must posit a social planner with preferences across the utilities of different agents populating the economy. I assume that the planner has weakly inequality-averse preferences across these agents' expected utilities at birth. These assumptions nest those of the dynastic, or Ramsey, model as a special case, while also allowing investigation of the shortcomings of that model.

The implications of this change are far-reaching. Introducing OLG alters the planner's problem in two distinct, conceptually opposing ways. First, it removes "Ricardian equivalence." In a representative-agent model, all possible plans for raising a fixed present value of revenue have the same income effects on household behavior and macroeconomic aggregates. However, in an OLG model, different plans will extract more or less income from a given generation and thereby have different income effects on that generation's behavior. The most direct implication is increased incentive for the government to run a balanced budget, as deficits result in accumulation of government debt, which drives up the interest rate, distorts capital accumulation and intertemporal consumption choices, and requires higher taxes in the future to make the elevated interest payments. This effect is stronger when interest rates respond more strongly to government debt—when the intertemporal elasticity of substitution (IES) is low, when the elasticity of substitution between capital and labor is low, and when access to international capital markets is limited, though the model in this paper is of a closed economy. Additionally, different timing of taxation leads to different labor supplies for the various generations and, in turn, different labor tax revenues. Specifically, delaying taxation makes current generations feel richer and work less, thereby giving the government less tax revenue now, but makes future generations feel poorer and work harder, thereby giving the government more tax revenue in the future. The net effect is of indeterminate sign, and is stronger when static income effects on labor supply are larger.

Second, as discussed above, OLG adds a redistribution motive to usual efficiency concerns, yielding an equity/efficiency tradeoff reminiscent of static, nonlinear optimal tax problems. This redistribution motive pushes toward smoother tax rates when government spending, and therefore taxes, are the primary reason for inequality across generations. When market forces lead to intergenerational inequality, this will push toward compensatory taxes that are higher on better-off generations. This effect is more pronounced when the planner is more inequality-averse.

The central analytical result of this paper shows how to incorporate these effects in forming optimal policy. In particular, the planner in an OLG economy considers all marginal effects of taxation: mechanical, substitution, and income. After summing these marginal effects at each point in time, the planner weights the result by the social marginal welfare weight of the generation paying the tax—the social value of an additional unit of consumption for that generation. He then smooths, or sets constant over time, that weighted sum—either state-by-state if markets are complete or in expectation if they are incomplete. By way of contrast, the social planner in a representative agent model need only smooth the marginal effect of linear taxation *above and beyond a hypothetical lump-sum tax*—the substitution effect, or "distortionary cost of taxation." This is not because the mechanical and income effects of the lump-sum component are irrelevant or not present, but rather because they are constant over time at the optimum—due to constant social marginal welfare weights and Ricardian equivalence—and thus are *necessarily* "smoothed" and cancel out.

I apply this framework to three distinct policy problems, chosen to highlight different aspects of optimal policy. First, I revisit the classic problem of optimal financing of wars, which highlights the partial tax smoothing properties of optimal policy in this framework. The loss of Ricardian equivalence pushes toward contemporaneous taxation to fund these wars, while the desire for intergenerational equity pushes back toward smooth taxes since nonconstant taxes are the proximate cause of inequality. Second, I consider an economy without required government spending but with shocks to productivity. Here, taxes should be higher during higher-productivity periods and interest payments on government debt (interpreted as Social Security payments) should be procyclical to improve intergenerational equality.[3] Finally, I incorporate trend productivity growth, which creates a persistent desire to redistribute from later generations to earlier ones. This is performed by increasing the average levels of government debt, interest rates, and therefore payments to retirees, which is funded by increasing the average labor tax rate—a system highly reminiscent of Social Security. Such systematic intergenerational redistribution is larger for more inequality-averse planners, who perceive more benefit from such redistribution and are therefore willing to accept higher distortionary costs from funding it.

As a brief preview of these numerical results, Figure 1.2 considers the optimal policy response to a completely unanticipated, single-period shock to government spending—an unanticipated war of known duration and size. The figure shows that in the Ramsey model, the optimal response is to fund the war entirely with debt and then permanently increase taxation to cover the ensuing interest payments, but never to repay the principal. Optimal policy in overlapping generations models, on the other hand, features substantial contemporaneous taxation to fund the war, with a return to the previous level of taxation in the long run. The rate of that reversion is inversely related to the degree of concavity in the social welfare function of the planner (parameterized by ζ). In the limit of a "Rawlsian" planner who cares only about the worst-off generation (an infinitely concave social welfare function), taxes never revert to their previous level, but instead remain high to fund compensatory transfers, in the form of large interest payments on debt, to the old. This figure will be presented and

[3] This pattern is reversed if income effects dominate.

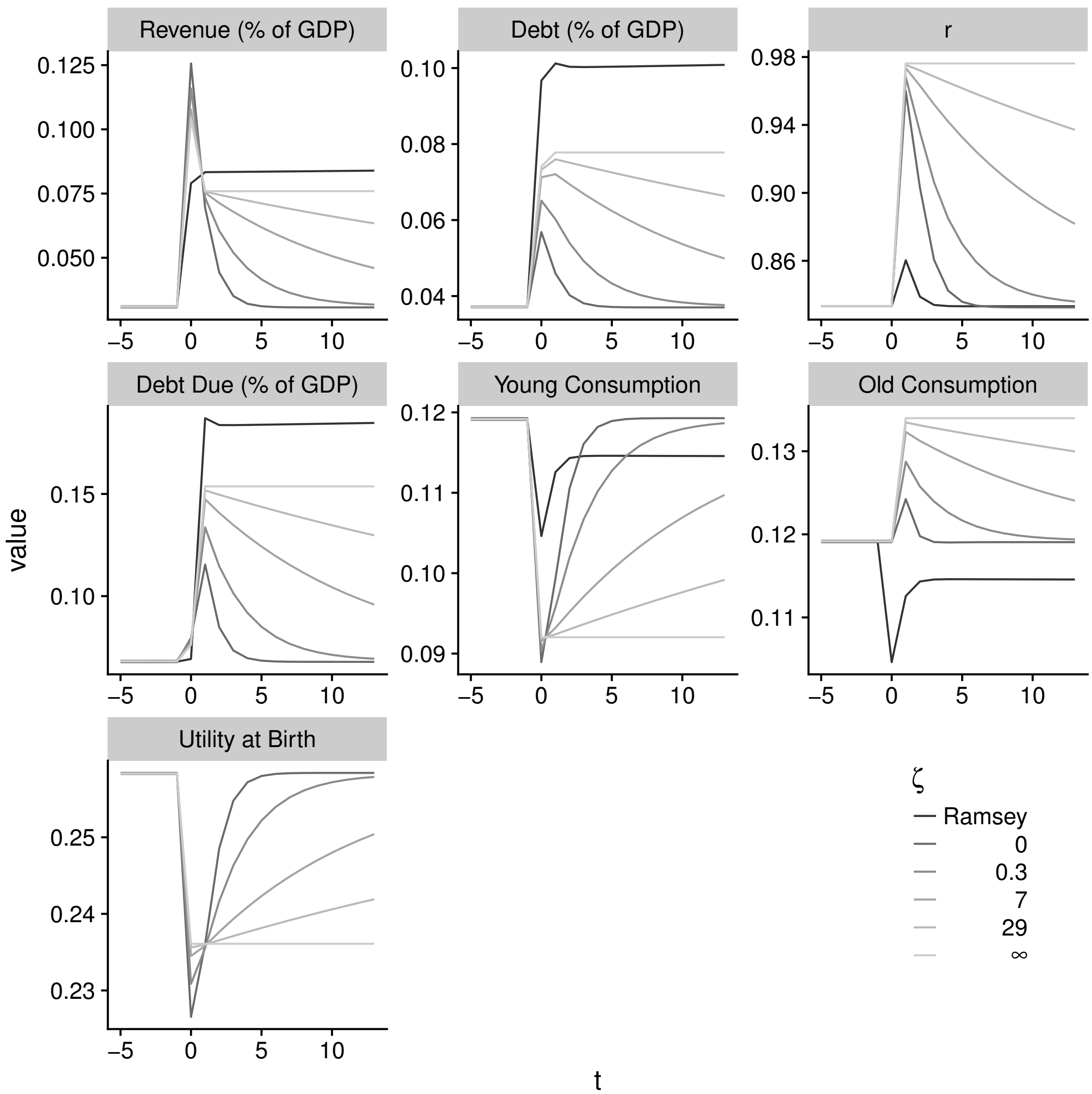


Figure 1.2: Optimal policy response to a completely unanticipated government spending shock equal in size to 10% of steady state GDP. These graphs show many economic variables each period, where period 0 is the period of the shock. Multiple models are considered: a Ramsey model, along with several OLG models featuring different planners with different levels of inequality-aversion, parameterized by ζ. All models assume isoelastic utility with an IES of 5.

analyzed again, with its assumptions more precisely specified, in Section 1.5; for now, merely consider how quantitatively and qualitatively different the optimal policies are for the representative-agent and OLG models.

1.1.1 Related Literature

This project relates to four major strands of economics literature.

Ramsey Taxation. The most directly related of these, referenced above, is commonly known as the Ramsey tax literature. This literature focuses on optimal linear taxation of a variety of goods in a model with a representative agent. Diamond and Mirrlees (1971) found this literature with consideration of the problem in a static context. Later, this problem was reformulated for a dynamic economy, with an eye toward aggregate uncertainty, by Barro (1979), who studies a reduced form model in which a planner seeks to minimize distortionary costs of taxation, which are assumed quadratic in the tax rate. This leads to a martingale property of tax rates. Additional papers in this strand consider richer models involving utility maximization and endogenous distortionary costs of taxation. Major findings include various versions of labor tax smoothing (Lucas and Stokey, 1983) and zero capital taxes (Judd, 1985; Chamley, 1986). A nice summary of this sort of problem can be found in Chari and Kehoe (1999).

Aiyagari et al. (2002) and, later, Farhi (2010) extend this literature by considering models with incomplete markets, in which the government does not have access to state-contingent debt to insure itself against adverse shocks. This leads to a smoothing of distortionary costs of taxation in expectation rather than state-by-state, and an employment of capital taxes[4] to hedge the government budget and manipulate interest rates. Aiyagari (1995) replaces aggregate uncertainty with idiosyncratic productivity

[4]if capital is present, as in Farhi (2010) but not Aiyagari et al. (2002)

shocks among infinitely-lived, *ex ante* homogeneous households and also finds that nonzero capital taxes are optimal.

Many papers analytically reconsider Ramsey taxation within the context of an OLG model (Atkinson and Sandmo, 1980; Escolano, 1992; Erosa and Gervais, 2002; Garriga, 2017). Other papers quantitatively investigate optimal policy in OLG models. For example, Conesa, Kitao and Krueger (2009) looks for optimal labor and capital taxes in a nonstochastic economy where policy must be constant and follow a given parametric form. While these papers might seem similar to the present one in concept, they differ in two critical ways. First, they do not consider aggregate uncertainty and so can only inform policy at a nonstochastic steady state or the transition thereto; they fail to provide guidance on the proper response to stochastic shocks to the government budget. Second, their focus is on capital taxes—especially whether they should be zero in steady state or transition; while I briefly treat capital taxes, my focus is primarily on the response of labor taxation to shocks. Finally, Weinzierl (2011) and Erosa and Gervais (2012) focus on how taxes should vary over the life cycle—a possibility I preclude in the present paper—in a Ramsey problem with idiosyncratic shocks but still no aggregate uncertainty.

Mirrlees Taxation. The other half of optimal taxation literature focuses on nonlinear, redistributive taxation and is named after the seminal contribution of Mirrlees (1971). This paper considers the optimal nonlinear income tax in a static model featuring heterogeneous agents with a continuum of earning abilities. Many recent papers have added realism to this stylized model by considering a dynamic, life-cycle model with gradually unfolding uncertainty. Farhi and Werning (2013) considers households receiving idiosyncratic shocks to their exogenous wage over time and how these shocks might be optimally insured through a fully general, dynamic taxation system. Stantcheva

(2016) adds endogenous human capital acquisition—and government policies thereto pertaining—to this model.

This literature relaxes the Ramsey assumptions of linear or affine taxation and representative households, but on the other hand assumes no aggregate uncertainty or general equilibrium effects; instead, it only requires that the government break even in expectation and assumes a constant interest rate and exogenous effective wages. Werning (2007) attempts a compromise by allowing aggregate uncertainty and (in parts of the paper) nonlinear taxation of heterogeneous households. However, this heterogeneity is perfectly persistent over time. That is, some households have higher earning ability than others, but after the beginning of the model, households face no idiosyncratic uncertainty.

The model in the present paper features nonidentical households and, therefore, a redistributive motive for the planner, while allowing for aggregate uncertainty. However, heterogeneity and idisoyncratic uncertainty are present only in the form of birth cohort; each birth cohort is perfectly homogeneous, and no idiosyncratic uncertainty exists after birth.

Intergenerational Risk Sharing. Third is a literature explicitly focusing on issues of overlapping generations and intergenerational risk sharing. Samuelson (1958) and Diamond (1965) lay the groundwork for the OLG model I use throughout the paper and derive certain results about optimal taxes, but in a context that allows achievement of a first-best outcome through lump-sum taxation. Piketty and Saez (2013) and Farhi and Werning (2010) consider optimal bequest or inheritance taxation in the context of heterogeneous earning ability; I abstract from bequest taxation in the present project, and my numerical simulations focus on calibrations without a bequest motive. Farhi et al. (2012) considers a government lacking access to a commitment technology in an overlapping generations framework, which leads to an incentive to redistribute capital *ex post* in a model featuring intra-cohort heterogeneity but no aggregate uncertainty.

A related literature considers intergenerational risk sharing from outside the context of optimal taxation. Green (1977) asks whether social insurance against uncertain population growth in an overlapping generations model could be designed in a Pareto-improving manner; he concludes that such is analytically possible, but not for reasonable values of the parameters. Ball and Mankiw (2007) consider the problem of intergenerational risk sharing not from the point of view of a utilitarian social planner, but rather by asking what the complete markets outcome would be if individuals could trade Arrow-Debreu securities behind a veil of ignorance. They then consider implementing that allocation using more conventional taxation tools. Other, even more abstract papers (Gordon and Varian, 1988) exist as well.

OLG Tax Incidence. Finally is the literature assessing the incidence and welfare effects of policy perturbation in large-scale, quantitative, OLG simulation models. Auerbach and Kotlikoff (1987), Kotlikoff, Smetters and Walliser (1999), and Altig et al. (2001) consider numerous policy reforms, either small or large, in a 55-generation OLG model with perfect forsight—that is, without uncertainty—and analyze the transition path. They then assess which demographic groups are better- and worse-off under the reform. Though these papers involve models that are much richer than the present paper's, they are fundamentally answering an incidence question rather than an optimal policy question, while also abstracting from uncertainty. Another set of similar papers directly concerns Social Security—either its optimality, or consideration of specific reforms—rather than broader policy questions. As an example of the former, Harenberg and Ludwig (2014) find that introduction of a pay-as-you-go ("PAYGO") Social Security system is optimal when there are interacting idiosyncratic and aggregate risks, in contrast to, among others, Krueger and Kubler (2006), which considers only aggregate risk. As an example of the latter, Feldstein (1998) covers the subject of privatizing Social Security—one oft-discussed reform—in great detail.

The remainder of the paper is ordered as follows: In Section 1.2, I describe the economy and formally state the problem faced by the social planner. In Section 1.3, I build intuition by characterizing optimal policy in two cases leading to extreme policies: quasilinear utility, and log-separable utility. Section 1.4 characterizes optimal policy more generally, in models with either recursively complete or incomplete markets. Section 1.5 applies the model numerically to three different policy problems to give concreteness to the discussion. Finally, robustness to addition of further generations is tested in Section 1.6, and Section 1.7 concludes. All proofs can be found in the Appendix.

1.2 Model

In this section, I describe the economy and the policies available to the government. Afterwards, I define a few pieces of notation that will be useful for subsequent discussion of optimal policy. The economy is a closed, neoclassical overlapping generations (OLG) economy with two generations, aggregate risk, discrete time, elastic labor supply, and capital. The nature of the overlapping generations component is designed to nest a traditional infinite-horizon model. There are two variants of the model: one with recursively complete markets, and one with incomplete markets; these variants differ only in the available assets and policy instruments.

1.2.1 Uncertainty

Aggregate risk is described by a discrete set of states $s_t \in \mathcal{S}$ and histories of those states $s^t = (s_0, s_1, ...s_{t-1}, s_t)$. The state of the world s_t evolves according to a Markov process with transition matrix $\boldsymbol{P}$. Exogenous, required government spending g and labor-augmenting productivity A are each functions of the state of the world s_t (not its history), which captures any uncertainty, and t, which captures any trend growth: $g(s^t) = g(s_t, t)$ and $A(s^t) = A(s_t, t)$. There is no idiosyncratic uncertainty.

1.2.2 Available Assets

The only asset in positive net supply is productive and risky capital. In addition, other assets in zero net supply exist as below, depending on the variant of the model.

Complete Markets At each history s^t, a market opens for a set of state-contingent assets delivering consumption at date $t+1$. Specifically, for each $s_{t+1} \in \mathcal{S}$, there exists an asset which delivers one unit of consumption at history (s^t, s_{t+1}) and costs $q(s^t, s_{t+1})$ units of consumption at history s^t.

Despite the name, this model does not feature truly complete markets. Individuals may not purchase insurance against the generation or state of the world into which they are born. This limits the ability of a social planner to efficiently distribute risk between generations and generates the fundamental economic problem this model is designed to analyze.

Incomplete Markets The only other asset is a one-period risk free bond. For each s^t, this bond costs one unit of consumption at history s^t and delivers $R^f(s^t)$ units of consumption at all $s^{t+1} \succeq s^t$.

1.2.3 Agents

The economy consists of three types of agents: households, firms, and the government.

Government. I abstract from commitment issues and focus on a government with access to a perfect commitment technology.

Complete Markets. The government has access to linear taxes on labor income $\tau^L(s^t)$ and capital income $\tau^K(s^t)$, as well as non-negative, lump-sum transfers to the young $T(s^t)$.[5] Capital income taxes may be state-contingent, and as a result it can be assumed without loss of generality that they are levied on gross capital income. The government also has access to the state-contingent asset market, allowing it to structure its portfolio of assets and debt in such a way as to provide insurance. Thus, the government's budget constraint at history s^t is

$$b(s^t) + g(s_t, t) + T(s^t) \leq \sum_{s^{t+1} \succeq s^t} q(s^{t+1}) b(s^{t+1}) + \tau^L(s^t) w(s^t) \ell(s^t) + \tau^K(s^t) R^K(s^t) k(s^{t-1}) \tag{1.1}$$

where b is the government's debt due, w is the pre-tax wage, ℓ is total labor supply, R^K is the gross return to capital, and k is the level of capital. Without loss of generality, I assume that state-contingent assets are untaxed.

Incomplete Markets. The government behaves similarly if markets are incomplete, with a few important changes. Most importantly, the government no longer has access to the state-contingent asset market; it can only trade the risk free bond. Second, capital taxes may no longer be state contingent but must be set one period in advance. Finally, lump-sum transfers may no longer be assumed to accrue only to the young,

[5] This is equivalent to assuming age-dependent lump-sum transfers $T^y(s^t)$ and $T^o(s^t)$; since markets are recursively complete, anticipated transfers to the old can simply be converted into their present value by the young.

since households cannot convert anticipated, but uncertain, old age transfers into their present value. Thus, the government's budget constraint becomes

$$R^f(s^{t-1})b(s^{t-1}) + g(s_t, t) + T^y(s^t) + T^o(s^t) \leq b(s^t) + \tau^L(s^t)w(s^t)\ell(s^t) + \tau^K(s^{t-1})R^K(s^t)k(s^{t-1}), \qquad (1.2)$$

where b is now the government's debt issued, and other variables are unaltered. To avoid an unrealistic outcome in which the government accumulates sufficient assets to pay for all expenditures with interest on those assets, I impose a lower bound $\underline{b}(t)$ (which should be thought of as nonpositive) on government debt. To seriously enforce the idea that government debt is risk free, I impose an upper bound $\bar{b}(t)$ as well. Similar constraints are imposed by Aiyagari et al. (2002) and Farhi (2010).

Households. A series of overlapping generations of households live for two periods each, and there is no population growth. I abstract from the desire for intra-cohort redistribution by assuming that all members of a cohort are identical. They inherit $z(s^t)$, provide labor $\ell(s^t)$, and consume $c^y(s^t)$ during youth. During old age they consume $c^o(s^{t+1})$ and leave a bequest

$$z(s^{t+1}) \geq \underline{z}(s_{t+1}, t+1). \qquad (1.3)$$

They face a market wage $w(s^t)$, a lump-sum transfer $T(s^t)$, and a vector of asset prices $\boldsymbol{q}(s^{t+1})$. A household born at history s^t ranks allocations recursively according to[6]

$$U(s^t) = u(c^y(s^t), \ell(s^t)) + \beta\mathbb{E}_t[u(c^o(s^{t+1}), 0)] + \delta\mathbb{E}_t[U(s^{t+1})]. \qquad (1.4)$$

[6]I have imposed the assumption that utility is time-separable for ease of exposition. The results could easily be extended to cover non-time-separable utility.

δ parameterizes the degree of altruism toward offspring, or bequest motive, while $\underline{z}(s_t, t)$ limits the extent to which households may pass on debt to their heirs.[7] $\delta = \beta$ and $\underline{z} = -\infty$ corresponds to the traditional infinite-horizon model, with each generation allowed to die in an unlimited amount of debt but caring equally about its offspring as its old-age self. $\delta = 0$ and $\underline{z} = 0$ corresponds to the stark OLG economy, with households that do not care about their offspring and are not allowed to die in debt.

If markets are complete, households face a single budget constraint in present value terms:

$$c^y(s^t) + \sum_{s^{t+1} \succeq s^t} q(s^{t+1})(c^o(s^{t+1}) + z(s^{t+1})) \leq T(s^t) + z(s^t) + (1 - \tau^L(s^t))w(s^t)\ell(s^t) \quad (1.5)$$

If markets are incomplete, households face period-by-period budget constraints:

$$c^y(s^t) + k(s^t) + b(s^t) \leq T^y(s^t) + z(s^t) + (1 - \tau^L(s^t))w(s^t)\ell(s^t) \quad (1.6)$$

$$\begin{aligned} c^o(s^{t+1}) + z(s^{t+1}) \leq T^o(s^{t+1}) + R^f(s^t)b(s^t) \\ + (1 - \tau^K(s^t)R^K(s^{t+1})k(s^t) \quad \forall s^{t+1} \succeq s^t \end{aligned} \quad (1.7)$$

Households thus maximize (1.4) subject to (1.5) and (1.3) if markets are recursively complete, or subject to (1.6), (1.7), and (1.3) if markets are incomplete.

Firms. Firms have a constant returns to scale production technology $F(k, \ell; A)$, including returned capital net of depreciation. Facing given prices $w(s^t)$ for wages and $R^K(s^t)$ for capital, firms simply statically maximize over k and ℓ

$$F(k, \ell; A) - w(s^t)\ell - R^K(s^t)k.$$

[7] It should be thought of as nonpositive, though I do not formally impose that assumption.

1.2.4 Equilibrium and Implementability

Equilibrium is standard, consisting of prices, policies, and quantities such that agents optimize subject to budget constraints and markets clear. The formal definition is given in Appendix A.1.

To this point, I have been deliberately vague about the issue of the first period policy problem. As mentioned earlier, I assume a government with access to perfect commitment. This is not an innocuous assumption. As is typical of similar optimal taxation problems, the optimal unrestricted policy is time-inconsistent; the government wishes to confiscate wealth initially, but promise to never do so in the future. Fortunately, government behavior in the initial period is not important to the analytical results to follow, nor to many of the numerical simulation results, assuming a sufficient "start up period" that is ignored. However, in certain of the numerical simulations, initial conditions remain important in perpetuity. To ensure more realistic numerical solutions, I will forbid capital confiscation in the first period by assuming a zero capital tax rate, which is either exactly or approximately the nonstochastic steady-state optimal capital tax rate.

The optimal policy problem will be formulated using the well-known "primal approach," in which I search directly for the optimal allocation and afterward derive the supporting policies. To do so, I must characterize which allocations are achievable for some set of policies.

Definition 1.2.1 (Implementable Allocation) *An allocation is implementable if there exists a policy equilibrium of which it is a part.*

The conditions that characterize implementable allocations are similar between the complete and incomplete markets models, but I will specify both sets of conditions fully, as these form the basis of the social planner's problems to follow.

Proposition 1.2.1 (Implementability Conditions—Complete Markets) *An allocation* $\{c^o(s^t), c^y(s^t), \ell(s^t), k(s^t), z(s^t)\}_{t\geq 0}$ *is implementable in the complete markets model iff it*

- *Satisfies the resource constraint at each* $s^t, t \geq 0$*:*

$$c^y(s^t) + c^o(s^t) + g(s_t, t) + k(s^t) \leq F(k(s^{t-1}), \ell(s^t); A(s_t, t)) \tag{1.8}$$

- *Satisfies the implementability condition at each* $s^t, t \geq 0$*:*

$$u_c^y(s^t)c^y(s^t) + u_\ell^y(s^t)\ell(s^t) + \beta\mathbb{E}_t[u_c^o(s^{t+1})(c^o(s^{t+1}) + z(s^{t+1}))] \geq u_c^y(s^t)z(s^t) \tag{1.9}$$

- *Satisfies the following constraint on the initial old:*

$$c^o(s_0) + z(s_0) \geq b(s_0) + F_K(k_-, \ell(s_0); A(s_0))k_- \tag{1.10}$$

- *Satisfies the optimal bequest conditions at each* $s^t, t \geq 0$*:*

$$\beta u_c^o(s^t) \geq \delta u_c^y(s^t) \tag{1.11a}$$

$$z(s^t) \geq \underline{z} \tag{1.11b}$$

$$(\beta u_c^o(s^t) - \delta u_c^y(s^t))(z(s^t) - \underline{z}) = 0 \tag{1.11c}$$

Proposition 1.2.2 (Implementability Conditions—Incomplete Markets) *An allocation* $\{c^o(s^t), c^y(s^t), \ell(s^t), k(s^t), z(s^t)\}_{t\geq 0}$ *is implementable in the incomplete markets model iff it, together with some sequence* $\{b(s^t)\}_{t\geq 0}$

- *Satisfies the resource constraint (1.8) at each* $s^t, t \geq 0$
- *Satisfies the implementability condition for the young at each* $s^t, t \geq 0$*:*

$$u_c^y(s^t)(c^y(s^t) + k(s^t) + b(s^t)) + u_\ell^y(s^t)\ell(s^t) \geq u_c^y(s^t)z(s^t) \tag{1.12}$$

and the implementability condition for the old at each $s^t, t \geq 1$:

$$c^o(s^t) + z(s^t) \geq \frac{u_c^y(s^{t-1})}{\beta \mathbb{E}_{t-1}[u_c^o(s^t)]} b(s^{t-1}) + \frac{F_K(s^t) u_c^y(s^{t-1})}{\beta \mathbb{E}_{t-1}[F_K(s^t) u_c^o(s^t)]} k(s^{t-1}) \tag{1.13}$$

- *Satisfies the following constraint on the initial old:*

$$c^o(s_0) + z(s_0) \geq b_{-1} + F_K(k_-, \ell(s_0); A(s_0)) k_- \tag{1.14}$$

- *Satisfies the optimal bequest conditions (1.11) at each $s^t, t \geq 0$*

- *Satisfies the debt limits at each $s^t, t \geq 0$:*

$$\underline{b}(t) \leq b(s^t) \leq \bar{b}(t) \tag{1.15}$$

1.2.5 Social Planner

There exists a social planner, who must choose among the implementable allocations. He ranks allocations according to

$$\Delta^{-1} W[u(c_-^y, \ell_-) + \beta u(c^o(s_0), 0)] + \mathbb{E}_0 \sum_{t=0}^{\infty} \Delta^t W[u(c^y(s^t), \ell(s^t)) + \beta \mathbb{E}_t u(c^o(s^{t+1}), 0)] \tag{1.16}$$

where $W(\bullet)$ is a social welfare function, usually assumed to be weakly concave, and c_-^y and ℓ_- are young consumption and labor from last period. A few comments about the social planner's objective function are in order. First, he discounts later generations relative to earlier ones according to a social discount factor Δ. This should not be seen as a strong political economy assumption, but instead simply as a requirement to ensure a solution. Second, he values only that utility derived directly by the households—not the altruism they feel toward their descendants. This makes sense given that the planner directly values those descendants. Third, the argument of the social welfare function is *ex ante* expected lifetime utility. This allows a planner to have redistributive preferences across generations and across cohorts born at different histories s^t at the same time t, but not across households born at the same history s^t but experiencing different shocks in old age s_{t+1}. While inserting realized *ex post* utility as the argument of $W(\bullet)$ would allow such, it also fails to respect individual preferences over risk; if $W(\bullet)$ were strictly concave and the argument were realized *ex post* utility, then the social planner would be more risk averse than households and likely choose a constrained Pareto inefficient allocation. This assumption—that the planner respects individual preferences over risk—plays an important role by reducing the planner's desire to insure individuals against shocks they may face in old age relative to a more naive view one might take. Finally, if $W(\bullet)$ is strictly concave, notice that the planner's objective is not time-separable. For example, if an existing cohort of households experienced very poor youths, the government may wish to compensate them in their old age—a motive that cannot be captured in a time-separable objective.

While this set of assumptions is not completely general in that it does not trace the entire Pareto frontier—one could instead consider a set of Pareto weights across all possible cohorts such that the weights sum to one—it allows a reasonable degree of generality while preserving a structure that lends itself to a recursive formulation and numerical solution.

The planner thus faces the following problems:

Problem 1.2.1 (Complete Markets Planning Problem) *The social planner maximizes (1.16) subject to (1.8), (1.9), (1.10), and (1.11).*

Problem 1.2.2 (Incomplete Markets Planning Problem) *The social planner maximizes (1.16) subject to (1.8), (1.12), (1.13), (1.14), (1.11), and (1.15).*

These optimization problems are not generally convex. As a result, first order conditions are necessary, but not sufficient, for an optimum.

1.2.6 Notation

Here, I introduce some additional notation I will use throughout my analysis of optimal policy in this model. First, I attach the following Lagrange multipliers to the constraints associated with the two planning problems:

- $\Delta^t \psi(s^t)$ to the resource constraint (1.8)
- $\Delta^t \mu(s^t)$ to the implementability condition for the generation born at s^t in the complete markets model (1.9)
- $\Delta^t \mu^y(s^t)$ to the implementability condition for the young at s^t in the incomplete markets model (1.12)
- $\Delta^{t-1} \beta \mu^o(s^t)$ to the implementability condition for the old at s^t in the incomplete markets model (1.13)

Second, since I will no longer discuss pre-tax wages and instead write them as the marginal product of labor, I reuse $w(s^t) \equiv W'(s^t)$ as the derivative of the social welfare function for the generation born at s^t with respect to expected lifetime utility, while $\tilde{w}(s^t) \equiv u^y_c(s^t) w(s^t)$ represents the social marginal welfare weight for the cohort born at s^t—the value to the planner of an extra dollar of consumption for that cohort.

The formulas discussed in the next section make frequent reference to the (marginal) distortionary cost of taxation;[8] this means the marginal dead weight loss associated

[8] Many economists—especially those focused on static economies—are more familiar with optimal tax expressions in terms of the tax itself, or wedges, often in the form $\frac{\tau}{1-\tau}$. If there are no wealth effects, there does exist a nice relationship between the distortionary cost of labor taxation and the labor wedge:

$$\varepsilon \frac{\tau^L}{1-\tau^L} = \frac{\tilde{\mu}}{1+\tilde{\mu}},$$

with raising an additional dollar of revenue through a particular tax; put another way, it is the cost above and beyond a lump-sum tax that mechanically[9] raises the same amount of revenue. At the optimum, this must be equal to the improvement in the planner's objective function that would occur by allowing a dollar of non-distortionary, lump-sum taxation, normalized by the improvement in the planner's objective function that would occur through an extra dollar of consumption for the generation being taxed. Therefore, distortionary costs can be written in terms of other variables as follows:

Definition 1.2.2 (Distortionary Cost of Labor Taxation) *The (marginal) distortionary cost of labor taxation on the generation born at s^t, denoted $\tilde{\mu}(s^t)$, is defined as*

$$\tilde{\mu}(s^t) \equiv \frac{\mu(s^t)}{w(s^t)} \tag{1.17}$$

$$\tilde{\mu}(s^t) \equiv \frac{\mu^y(s^t)}{w(s^t)} \tag{1.18}$$

for the complete and incomplete markets cases, respectively.

Definition 1.2.3 (Distortionary Cost of *Ex Post* Capital Taxation) *The distortionary cost of* ex post *capital taxation on the generation dying at s^{t+1}, denoted $\hat{\mu}(s^{t+1})$, is defined as*

$$\hat{\mu}(s^{t+1}) \equiv \frac{\mu(s^t)}{w(s^t)} \tag{1.19}$$

$$\hat{\mu}(s^{t+1}) \equiv \frac{\mu^o(s^{t+1})}{w(s^t)u_c^o(s^{t+1})} \tag{1.20}$$

---

where ε is the elasticity of labor supply with respect to the net-of-tax wage. However, for more complex cases, such a simple relationship doesn't exist, and so the below expressions involving $\tilde{\mu}$ cannot be replaced with simple expressions involving τ^L. Additionally, $\tilde{\mu}$ has a much stronger intuitive meaning within the context of tax smoothing. Thus, I will work with $\tilde{\mu}$ in analytical expressions, though in numerical simulations I will give τ^L for concreteness.

[9] without accounting for behavioral response

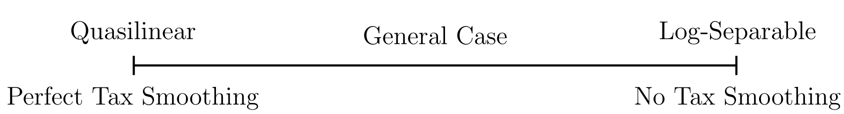
for the complete and incomplete markets cases, respectively.

The latter definition is less intuitive than the former, and warrants further discussion. First, in the case of complete markets, notice that the cost of capital taxation at any $s^{t+1} \succeq s^t$ is always the same as the cost of labor taxation at s^t. This is because at the optimum, revenue must be extracted from the generation born at s^t in an efficient way, or the resulting allocation will be constrained inefficient. Thus, the distortionary cost of all taxes applied to the same generation must be equal. Meanwhile, in the incomplete markets model, *ex post* capital taxes are not allowed; the capital tax rate must be set one period in advance. Nonetheless, one can consider the cost of raising such a tax if it were allowed; that is what $\hat{\mu}(s^{t+1})$ captures, and it matches the marginal value of allowing a small lump-sum tax on the old in only that particular state of the world. Since each such instrument is not actually available, the distortionary costs of all such instruments need not equal each other, or of labor taxation in the previous period.

Finally, from this point forward, I reduce dependence on s^t to a subscript t for all variables when it does not lead to a substantial loss in clarity.

1.3 Two Polar Cases

Before characterizing optimal policy for the general case, I build intuition for the main results by discussing optimal policy in two extreme cases: quasilinear utility and a utilitarian planner; and log-separable utility, a utilitarian planner, and no capital.



We will see that in the former case, the usual, perfect tax smoothing results obtain, while in the latter, policy is formed entirely period-by-period, with no tax smoothing or intertemporal concerns whatsoever. In both cases, I will focus on a pure OLG model—no bequest motive ($\delta = 0$), and a minimum bequest of 0 ($\underline{z}(s_t, t) = 0$).

1.3.1 Quasilinear Utility, Utilitarian Planner, $\beta = \Delta$

Proposition 1.3.1 *Assume that period utility is quasi-linear:*

$$u(c, \ell) = c - v(\ell),$$

where $v(\bullet)$ is increasing and convex, with negative consumption permitted. Further assume the planner is utilitarian with $\beta = \Delta$. If markets are recursviely complete, then for $t \geq 0$, the distortionary cost of labor taxation is constant

$$\tilde{\mu}_t = \tilde{\mu}_{t+1}$$

and ex ante capital taxes are zero

$$\mathbb{E}_t[\tau^K_{t+1} F_{K,t+1}] = 0.$$

If markets are incomplete, then the distortionary cost of labor taxation is a martingale away from debt limits

$$\tilde{\mu}_t = \mathbb{E}_t[\tilde{\mu}_{t+1}] + \nu^u_t - \nu^l_t$$

and capital taxes satisfy

$$\frac{\tau^K_t}{1 - \tau^K_t} = \frac{\mathrm{Cov}_t[\tilde{\mu}_{t+1}, k_t F_{KK,t+1}] - \frac{\mathbb{E}_t F_{KK,t+1}}{\mathbb{E}_t F_{K,t+1}} \mathrm{Cov}_t[\tilde{\mu}_{t+1}, k_t F_{K,t+1}]}{\mathbb{E}_t[(1 + \tilde{\mu}_{t+1}) F_{K,t+1}]}.$$

All of these results except the last are highly familiar elements of optimal tax literature; the last is equivalent to that in Farhi (2010), which serves as the benchmark for optimal taxation with incomplete markets. The standard results are preserved because the two effects that overlapping generations introduce are both eliminated in this specification. First, the planner has no redistributive preferences—all generations have the same social marginal welfare weight, since the planner is utilitarian and there is no deminishing marginal utility—so policy is constructed to maximize efficiency, as in a classic dynastic model. Second, quasilinear utility carries no income effects, thereby eliminating any difference in household behavior between OLG and dynastic versions of the model. With these two new effects nullified, the model resembles the classic Ramsey model.

1.3.2 Log-Separable Utility, No Capital, Utilitarian Planner, $\beta = \Delta$

Proposition 1.3.2 *Assume that period utility is log-separable:*

$$u(c, l) = \log c - v(\ell),$$

where $v(\bullet)$ is increasing and convex. Further assume the planner is utilitarian with $\beta = \Delta$, and there is no capital:

$$F(k, \ell; A) = A\ell.$$

If markets are recursively complete, then for $t \geq 1$, the distortionary cost of taxation (as well as the allocation and tax rate) depends only on the current state, and not on the stochastic process for government spending or initial conditions, so long as lump-sum transfers are not used.

The reasons for this stark result, which states that standard tax smoothing has no role in this economy,[10] are twofold. First, with log-separable utility and a single working

[10]For emphasis, there is nothing special about this setup for an infinite-horizon economy. Perfect

period per cohort, income and substitution effects of labor taxation cancel out, leaving labor effort unaffected. This means the government can set whatever labor tax rate it sees fit without any consequences for GDP. It is worth emphasizing that this does *not* mean that labor taxation is without efficiency costs; the first best would involve an *increase* in labor effort during a war. This serves to highlight a key point made in the more general analysis: With overlapping generations, income and mechanical effects matter—not just the standard substitution effects.

Second, log-separable utility has an intertemporal elasticity of substitution (IES) of 1. Combined with the availability of complete markets, this means that old consumpiton at s^{t+1} can be "chosen" by the planner, completely independently of the allocation at s^t as well as the allocations at other successors of s^t. It is this latter property—the fact that old consumption at s^{t+1} can be set independently of allocations at other possible sucessors of s^t—that the incomplete markets model lacks and, therefore, prevents it from having this extreme property.

These two properties, combined with the lack of capital, give the planner a completely time-separable problem.[11]

---

tax smoothing would apply as usual.

[11]It is worth noting that the sense in which optimal policy depends only on the present period here is quite different from a similarly-worded finding in Lucas and Stokey (1983). In that paper, optimal policy depends on the initial government budget position and the Markov process for government spending (through the sufficient statistic of the multiplier on the implementability condition) as well as the current government spending shock. This means that within any simulation of the model, optimal policy will feature the same tax rates in all periods that share the same government spending shock, but not across simulations that feature different initial conditions or different Markov processes. In constrast, optimal policy in the present model depends only on the current government spending shock; neither initial conditions nor the properties of the Markov process are relevant. Moreover, in Lucas and Stokey (1983), the distortionary cost of taxation is constant over time—the usual tax smoothing result—and depends on the initial conditions and the Markov process for government spending. In contrast, the distortionary cost of taxation here is *not* constant, and also depends only on the current government spending shock, and neither the initial conditions nor the properties of the Markov process.

1.4 Properties of Optimal Policy in the General Case

Having discussed two special, polar cases—one involving perfect, traditional tax smoothing, and another in which current policy depends only on the present shock—I now address the properties of optimal policy in the general case. It features elements of both extreme cases presented last section: tax smoothing for usual efficiency reasons, and taxes that depend on the current shock due to income effects and the loss of Ricardian equivalence; meanwhile, the redistributive motive can push toward smoother or less smooth taxes depending on the nature of shocks. At the optimum, the planner chooses the best possible tradeoff among these goals in a way this section will make precise. I first discuss the mathematically simpler case of recursively complete markets before considering the slightly messier case of incomplete markets, though the cases are conceptually similar.

1.4.1 Recursively Complete Markets

Labor Taxes

I begin by characterizing optimal labor taxes. When markets are complete, the government can reform its policies in any way that is budget neutral in present value, raising revenue at any combination of histories it sees fit. A candidate policy is optimal only if no such reform improves welfare. These results focus on the neutrality of a small reform in which the government increases labor taxes at history s^t and uses the proceeds to reduce labor taxes at a particular $s^{t+1} \succeq s^t$. Since these two histories are temporally adjacent, a "government Euler equation" with respect to a particular state-contingent asset emerges. The nature of such an Euler equation depends strongly on whether the bequest requirement binds.

Proposition 1.4.1 *If markets are complete and the bequest requirement does not bind at s^{t+1}, then the distortionary cost of labor taxation satisfies*

$$q(s^{t+1})\tilde{w}_t\tilde{\mu}_t = Pr_t(s^{t+1})\Delta\tilde{w}(s^{t+1})\tilde{\mu}(s^{t+1}). \tag{1.21}$$

This equation resembles the standard tax smoothing result. It differs only to the extent that

$$q(s^{t+1})\tilde{w}_t \neq Pr_t(s^{t+1})\Delta\tilde{w}(s^{t+1}).$$

Expanding q and noting that bequests are set optimally when the bequest requirement does not bind shows this is equivalent to the relation

$$\delta w_t \neq \Delta w(s^{t+1}).$$

That is, at histories at which the bequest requirement does not bind, the planner deviates from tax smoothing only to the extent that the relative social weighting of the two adjacent generations differs from the relative private weighting. If the planner is utilitarian and $\delta = \Delta$, we recover the Ramsey tax smoothing result. These possibly differing weights embody the essence of the redistributive motive—one of the two effects introduced by OLG. The other—the loss of Ricardian equivalence—does not factor in this relation because households are, in fact, locally Ricardian; since bequests are active, anticipated taxes at s^{t+1} are felt proportionately by the generation born at s^t and lead to an adjustment of their consumption path.

Proposition 1.4.2 *If markets are complete and the bequest requirement binds at s^{t+1},* *then the distortionary cost of labor taxation satisfies*

$$q_{t+1}\,\tilde{w}_t \overbrace{\left\{\underbrace{1}_{\substack{\text{Mech.}\\ \text{Effect}}} + \tilde{\mu}_t \left[\underbrace{1}_{\substack{\text{Subst.}\\ \text{Effect}}} - \underbrace{\sigma^o_{t+1}\left(1 + \frac{\underline{z}_{t+1}}{c^o_{t+1}}\right)}_{\substack{\text{Dynamic Income}\\ \text{Effect}}}\right]\right\}}^{\text{M.C. of labor taxation at } s^t} = Pr(s^{t+1})\Delta\,\tilde{w}_{t+1} \overbrace{\left\{\underbrace{1}_{\substack{\text{Mech.}\\ \text{Effect}}} + \tilde{\mu}_{t+1}\eta_{t+1}\right\}}^{\text{M.C. of labor taxation at } s^{t+1}} \tag{1.22}$$

where

$$\sigma^j_t = -\frac{u^j_{cc,t}c^j_t}{u^j_{c,t}} = \frac{1}{IES^j_t}$$

and

$$\begin{aligned}
\eta_{t+1} =& 1 - \sigma^y_{t+1}\left(1 - \frac{\underline{z}_{t+1}}{c^y_{t+1}}\right) + \frac{u^y_{c\ell,t+1}}{u^y_{c,t+1}}\ell_{t+1} \\
=& \underbrace{1}_{\substack{\text{Subst.}\\ \text{Effect}}} + \underbrace{\ell_{t+1}\left(\frac{u^y_{c\ell,t+1}}{u^y_{c,t+1}} + F_{L,t+1}(1-\tau^L_{t+1})\frac{u^y_{cc,t+1}}{u^y_{c,t+1}}\right)}_{\text{Static Income Effect}} \\
& + \underbrace{\sigma^y_{t+1}\frac{F_{L,t+1}(1-\tau^L_{t+1})\ell_{t+1} + \underline{z}_{t+1} - c^y_{t+1}}{c^y_{t+1}}}_{\text{Dynamic Income Effect}}.
\end{aligned}$$

This expression considers periods when the bequest requirement does bind—all periods in a strict OLG model. The left hand side represents the marginal cost of labor taxation per dollar raised at history s^t, and the right hand side represents the marginal cost at history s^{t+1}.

The former has three components: the mechanical effect of the tax[12], the distortionary cost or substitution effect of the tax[13], and the dynamic income effect of the tax, which acts through changing asset prices and is less familiar in optimal tax analysis.

[12]Increased taxes reduce post-tax income and, therefore, utility.

[13]Increased taxes reduce post-tax wages and, therefore, the incentive to work, costing the government some tax revenue.

Specifically, the reform under consideration reduces old consumption at s^{t+1} but does not alter young consumption at s^t.[14] Thus, $q(s^{t+1})$ must rise, which makes households born at s^t feel poorer, work harder, and give the government additional tax revenue, which offsets the distortionary cost of taxation. This dynamic income effect is proportional to the inverse of the IES, which governs how much asset prices respond to changes in consumption trajectories.

The right hand side has four distinct components. First is the mechanical effect, and η_{t+1} encompasses the other three: the standard distortionary cost or substitution effect, the standard (static) income effect which is countervailing and proportional to ℓ_{t+1}[15], and a dynamic income effect similar to the one above. Under this perturbation, consumption at s^{t+1} falls but consumption at s^{t+2} stays the same for all $s^{t+2} \succeq s^{t+1}$. Thus, the price of consumption at all histories $s^{t+2} \succeq s^{t+1}$ falls. This makes the household born at s^{t+1} feel richer in present value terms, decrease labor supply, and partially offset the standard income effect/exacerbate the substitution effect. This effect is proportional to saving, and again proportional to the inverse of the IES.

The final ingredient is the social marginal welfare weight, which captures the idea that the planner may care more about costs imposed on one generation than another. The sum of all effects on each cohort, weighted by this social marginal welfare weight, must be equal at the optimum in present value terms. Otherwise, the planner could benefit from a small reform.

This central result makes two conceptual alterations to the standard tax smoothing

---

[14] To see this, notice that since the government has directed all of the new revenue, ϵ, in this reform to reduction in taxes at s^{t+1}, then $b(s^{t+1})$ drops by $\frac{\epsilon}{q(s^{t+1})}$ with the government budget position unaltered at all other histories. Market clearing requires that $c^o(s^{t+1}) + z(s^{t+1})$ must also reduced by $\frac{\epsilon}{q(s^{t+1})}$, which precisely accounts for the lost income from the tax, leaving young consumption at s^t unchanged.

[15] When consumption drops, households feel poorer, and therefore work harder. This gives the government additional tax revenue, which offsets the distortionary cost of taxation.

result of the Ramsey model—or the quasilinear case from last section—directly corresponding to the two new considerations introduced by the OLG framework. First, it takes into account income effects, which reduce the distortionary costs of taxation. These could be ignored in the Ramsey model, or if bequest limits do not bind, since they do not depend on tax policy due to Ricardian equivalence, but cannot be ignored in an OLG model if bequest requirements bind. Second, it weights by the social marginal welfare weight, which captures the desire to redistribute. Social marginal welfare weights will generically be nonconstant even if the planner is utilitarian, since if the bequest constraint binds, marginal utilities of consumption will be unequal across generations.[16]

These two alterations have directly opposing effects. A larger departure from Ricardian equivalence, or stronger income effects, leads to a smaller (or negative) weight on the distortionary cost of taxation, thereby encouraging higher taxes on less well off generations. Viewed from the government budget perspective, the "less Ricardian" are households, the larger is the distortion in interest rates associated with government saving and borrowing, and the more positively correlated are labor supply and taxes due to income effects, each of which pushes toward balanced government budgets. On the other hand, a more concave social welfare function means a stronger redistributive motive, lower taxes on less well-off generations, and therefore counteracts the loss of Ricardian equivalence. This effect can push toward smoother taxes if the main source of inequality across generations is different tax rates; it can push toward less smooth taxes if generations' pre-tax wages are highly unequal.

[16]This weighting also necessitates adding back in mechanical effects, which are implicitly subtracted from both sides in the Ramsey case where social marginal welfare weights are constant.

Capital Taxes

One could instead consider a reform that increases labor taxes at history s^t and uses the proceeds to reduce capital taxes by an equal amount at all $s^{t+1} \succeq s^t$. This reform should be welfare-neutral at the optimum, and analyzing it reveals properties of optimal capital taxes. As the zero capital tax result of Chamley (1986) and Judd (1985) is quite familiar, I focus here on the stark OLG case.

Proposition 1.4.3 *If markets are complete and bequest limits bind at zero ($\underline{z}(s_t, t) = 0$, and $\delta = 0$), then optimal capital taxes satisfy*

$$\sum_{s_{t+1}} q_{t+1} F_{K,t+1} \tau^K_{t+1} = \frac{\tilde{\mu}_t}{1 + \tilde{\mu}_t} \left\{ \eta_t - 1 + \sum_{s_{t+1}} q_{t+1} F_{K,t+1} \sigma^o_{t+1} \right\}$$

Recall that $\eta_t - 1$ includes static income effects (which are negative) and dynamic income effects (which are positive). Thus, this proposition says that ex ante capital taxes are more positive when any of the following is true:

- Static income effects are smaller. Raising capital taxes causes households to save less and consume more in youth. This causes households to supply less labor and therefore pay less labor income tax, which hurts the government budget. This effect is muted when static income effects are smaller.

- Dynamic income effects are larger. Raising capital taxes means that consumption in old age is more expensive. This makes households feel poorer, which causes them to supply less labor and therefore pay more labor income tax, which helps the government budget.

- The IES is smaller. The larger the IES, the greater the distortion of capital accumulation that comes from an increase in capital taxes.

The magnitude of the capital taxes—positive or negative—increases when the labor tax rate, and therefore the marginal distortionary cost of labor taxation, is higher, because the effects on the government budget are amplified. There is a classic case in which these three effects offset and ex ante capital taxes are precisely zero:

Corollary 1.4.4 *If preferences exhibit constant relative risk aversion σ over consumption, and are separable between labor and consumption, then ex ante capital taxes between t and $t+1$ are zero if bequest limits bind at zero.*

This draws upon the static Ramsey tax intuition: Goods should be taxed at rates inversely related to their elasticities of demand. If utility is homothetic over consumption and separable from labor, then all states' consumption should be taxed equally, which means that capital should not be taxed.

1.4.2 Incomplete Markets

Moving to incomplete markets does not change any fundamental intuition. It merely adds complexity in that no feasible perturbation affects just two contiguous histories; generically, any perturbation affecting history $s^{t+1} \succeq s^t$ will affect $(s^t, s_{t+1}) \, \forall s_{t+1} \in \mathcal{S}$. Thus, expected returns to *both* assets (risk free debt and capital) will generally be affected. These effects must be considered when weighing a policy perturbation.

Labor Taxes

I will present two expressions that enforce the neutrality of a small reform in which the government raises labor taxes at history s^t, which it uses to reduce its risk free debt and, in turn, reduce labor taxes collected at all $s^{t+1} \succeq s^t$ by an equal amount. At the optimum, this reform must be welfare-neutral. The expressions differ, as those in the last section did, depending on whether the bequest constraint binds at s^{t+1}. I begin with an expression analogous to (1.21):

Proposition 1.4.5 *If markets are incomplete and the bequest requirement does not bind for any $s^{t+1} \succeq s^t$, then the distortionary cost of labor taxation satisfies*

$$\tilde{w}_t \tilde{\mu}_t = \Delta R_t^f \mathbb{E}_t[\tilde{w}_{t+1} \tilde{\mu}_{t+1}] + \nu_t^u - \nu_t^l \tag{1.23}$$

where R_t^f is the gross risk free rate between t and $t+1$.

This expression sums (1.21) across all $s^{t+1} \succeq s^t$, with an adjustment for the possibility of binding government debt limits, since market incompleteness prevents the government from performing a reform involving only one $s^{t+1} \succeq s^t$. If the government has access to credit, the intuition is identical to (1.21): The planner trades off the gap between private intergenerational weighting and social intergenerational weighting against the ratio of distortionary costs, taking into account any covariance between the two.

As with complete markets, this expression differs from its Ramsey equivalent—that distortionary costs of taxation are a risk-adjusted martingale—only to the extent that $\delta w_t \neq \Delta w_{t+1}$. Thus, the only alteration induced by OLG is the motive for redistribution; when the request requirement does not bind, households are indeed locally Ricardian.

Next, consider an expression analogous to (1.22):

Proposition 1.4.6 *If markets are incomplete, bequest limits bind at zero, and govern-*

ment debt limits do not bind, then the distortionary cost of labor taxation satisfies

$$\tilde{w}_t \left\{ 1 + \tilde{\mu}_t \left(1 - \underbrace{\epsilon_{b_t}^{R_t^f}}_{\substack{\text{Updated Dynamic}\\ \text{Income Effect}}} \right) + \underbrace{\frac{\partial \tau_t^K}{\partial b_t}}_{\substack{\text{Size of Capital Tax}\\ \text{Windfall}}} \sum_{s_{t+1}} \left[\underbrace{q_{t+1} F_{K,t+1} k_t \hat{\mu}_{t+1}}_{\substack{\text{Benefit of Capital Tax}\\ \text{Windfall}}} \right] \right\}$$
$$= \Delta R_t^f \mathbb{E}_t \left\{ \tilde{w}_{t+1} \left[1 + \tilde{\mu}_{t+1} \eta_{t+1} + \underbrace{\frac{\sigma_{t+1}^y (1 - \tau_{t+1}^K)}{c_{t+1}^y}}_{\substack{\text{Size of Capital Tax}\\ \text{Reduction}}} \left(\begin{array}{c} \sum_{s_{t+2}} \left[\underbrace{q_{t+2} F_{K,t+2} k_{t+1} \hat{\mu}_{t+2}}_{\substack{\text{Cost of Capital Tax}\\ \text{Reduction}}} \right] \\ - \underbrace{\frac{\tilde{\mu}_{t+1} k_{t+1}}{1 - \tau_{t+1}^K}}_{\substack{\text{Relative to Previous}\\ \text{Dynamic Income Effect}}} \end{array} \right) \right] \right\} \tag{1.24}$$

where q_{t+1} is the price of a state-contingent asset delivering consumption at s^{t+1}, were such an asset available, and

$$\epsilon_{b_t}^{R_t^f} = \frac{\partial R_t^f}{\partial b_t} \frac{b_t}{R_t^f} = -R_t^f b_t \frac{\mathbb{E}_t u_{cc,t+1}^o}{\mathbb{E}_t u_{c,t+1}^o}.$$

Relative to summing (1.22) over all $s^{t+1} \succeq s^t$, this expression involves several alterations, highlighted with underbraces, though the overall intuition is unaltered. First, the left hand side replaces σ_{t+1}^o with $\epsilon_{b_t}^{R_t^f}$, which plays the same role—it captures the income effect of a tax increase at date t, with all of the associated drop in consumption occurring at date $t+1$.[17] When the government raises taxes at date t, fewer bonds are issued, necessitating a drop in the interest rate, which makes households born at t feel poorer. This in turn makes them work harder, offsetting some of the distortionary effect—a

[17]To see that they are equivalent, notice that since σ_{t+1}^o is the inverse of the intertemporal elasticity of substitution for the old at s^{t+1}, it is also the percentage that the gross return on the asset paying off at s^{t+1} changes when $c^o(s^{t+1})$—i.e., the amount of that asset issued—increases by one percent, holding fixed $c^y(s^t)$. Similarly, $\epsilon_{b_t}^{R_t^f}$ is the percentage that the risk free gross return changes when b_t increases by one percent (and therefore all $c^o(s^{t+1})$ increases by $1\% \cdot b_t R_t^f$).

dynamic income effect.[18]

Second, $\frac{\partial \tau_t^K}{\partial b_t} \sum_{s_{t+1}} [q_{t+1} F_{K,t+1} k_t \hat{\mu}_{t+1}]$ is added to the left hand side. This reflects the fact, just discussed, that a perturbation in which tax revenue increases at t and falls by R_t^f at $t+1$ necessitates a drop in the interest rate. But in equilibrium, the after-tax return on capital must drop as well. This could occur through increase in capital accumulation, but this fails to keep the perturbation confined to periods t and $t+1$. Instead, equilibrium can be achieved via an increase in capital taxes equal to $-\frac{\partial \tau_t^K}{\partial b_t}$,[19] which in turn leads to an increase in capital tax collections at date $t+1$ by $-\frac{\partial \tau_t^K}{\partial b_t} F_{K,t+1} k_t$. These increased collections act as free lump-sum taxes on the old at date $t+1$, since they were enacted precisely to *prevent* behavior from changing; therefore, they represent a gain for the government budget worth $q_{t+1}\hat{\mu}_{t+1}$ more than their cost to households, by definition of $\hat{\mu}$. Similarly, $\frac{\sigma_t^y (1-\tau_t^K)}{c_t^y} \sum_{s_{t+1}} [q_{t+1} F_{K,t+1} k_t \hat{\mu}_{t+1}]$ is added the right hand side. This term is of the opposite sign since capital taxes *drop* between $t+1$ and $t+2$ which hurts the government budget position, whereas they *rise* between t and $t+1$ which helps the government budget position. Finally, $\sigma_{t+1}^y \tilde{\mu}_{t+1} \frac{k_{t+1}}{c_{t+1}^y}$ is subtracted from the right hand side. This is to avoid double-counting the loss that occurs from the drop in capital taxes.[20]

Though this expression appears far more complicated than (1.22), it captures the

[18]The right hand side needs no similar adjustment from the summed (1.22), because the reduction in c_{t+1}^y has the same effect on all of the state-contingent asset returns as it does on the risk free rate.

[19]recall that b_t drops if taxes are increased at date t

[20]To see this, note that the addition of $\frac{\sigma_t^y (1-\tau_t^K)}{c_t^y} \sum_{s_{t+1}} [q_{t+1} F_{K,t+1} k_t \hat{\mu}_{t+1}]$ to the right hand side, captures the entire effect of a change in capital taxes, holding fixed capital accumulation and, therefore, pre-tax capital returns. However, recall η_{t+1} includes a "dynamic income effect" that captures the fact that when young consumption at date $t+1$ drops, interest rates rise, making households feel richer and work less, hurting the government budget position. η_{t+1} applies this effect to all of saving—both risk free bonds and capital—while the second term discussed above already accounts for this effect on capital. Thus, this final term is subtracted to counteract that. This correction isn't necessary on the left hand side because $\epsilon_{b_t}^{R_t^f}$, as an elasticity, only is proportional to b_t and not all of saving $b_t + k_t$.

same intuition: If bequest limits bind, then optimal labor taxes must be set to balance mechanical, substitution, and income effects, with the latter further broken down into static and dynamic income effects. However, whereas the right hand sides of (1.22) for two different $s^{t+1} \succeq s^t$ can be set equal to create a tax smoothing result across different states of the world at $t+1$, market incompleteness forbids such inter-state smoothing here, forcing the planner to focus exclusively on intertemporal smoothing.

Capital Taxes

As with complete markets, one can consider a policy perturbation wherein labor taxes at date t are increased and capital taxes between dates t and $t+1$ are cut. At the optimum, this must be welfare-neutral, and gives rise to the following result.

Proposition 1.4.7 *If markets are incomplete, bequest limits bind at zero, and government debt limits do not bind, then capital taxes satisfy*

$$\frac{\tau_t^K}{1-\tau_t^K} = \frac{\left\{\begin{array}{c} \mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]\tilde{\mu}_t\left[\left(\eta_t - 1 - \sigma_t^y \frac{k_t}{c_t^y}\right) + \frac{1}{1-\tau_t^K}\epsilon_{k_t}^{R_t^f}\frac{b_t}{k_t}\right] \\ +\frac{1}{1-\tau_t^K}\mathbb{E}_t\left[F_{K,t+1}u^o_{c,t+1}\epsilon_{k_t}^{1-\tau_t^K}\hat{\mu}_{t+1}\right] \\ +\mathrm{Cov}_t[\hat{\mu}_{t+1}, k_t F_{KK,t+1}u^o_{c,t+1}] \\ -\frac{\mathbb{E}_t[F_{KK,t+1}u^o_{c,t+1}]}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]}\mathrm{Cov}_t[\hat{\mu}_{t+1}, k_t F_{K,t+1}u^o_{c,t+1}] \end{array}\right\}}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}(1+\hat{\mu}_{t+1})]} \tag{1.25}$$

This equation is related to the benchmark result in Farhi (2010), but noticeably different. The last row of the numerator is perfectly analogous to that paper's "hedging term," and encourages higher capital taxation when the return to capital is correlated with the government's desire for funds, tempered by any adjustement in capital accumulation. That paper's other term was labeled the "intertemporal term," and "reflects the possibility of manipulating interest rates." However, the essence of that term was that capital taxes should be raised if there is an adverse shock at date t, and not if there isn't; this leads to higher interest rates for the government if there is an adverse shock, but lower interest rates if not. Two complexities make this term not applicable to the present model. First, those two different levels of capital taxes fall on different cohorts, which necessitates consideration of welfare weights or a different policy perturbation. Second, while such policy plans may be "anticipated" by households the period before, they are not *internalized*, since they will only affect the newly born generation.[21] Instead of this intertemporal term, the first row contains the income effects of such a policy perturbation, which does not need to be considered in a representative-agent model.

1.4.3 Summary

In this section, we saw that moving from the Ramsey model to an OLG model introduces additional effects the planner must consider when setting optimal taxes. First, taxes have (nonconstant) static and dynamic income effects due to loss of Ricardian equivalence. The former embodies the classic offsetting income effect in labor tax analysis, while the latter captures the fact that altering the policy path alters asset prices. Second, redistribution motives generally exist—even when the planner if utilitarian if there is diminishing marginal utility of consumption. Though these effects are present in both complete and incomplete markets models, the expressions satisfied by optimal policies are more complicated in the latter case, since attention must be given to the which perturbations are possible.

[21] This is, of course, an implication of the extreme assumption that households live for only two periods, and make meaningful economic choices in only one. Nonetheless, the intuition would remain that the effects of anticipated capital tax policy in the future would not be fully internalized by generations alive today, which would mitigate the value of the intertemporal term.

1.5 Three Applications to Policy Questions

Having discussed in detail the analytic properties of optimal policy in models with overlapping generations, I now give more concrete understanding of how governments ought to conduct policy in such models through numerical application to three important, conceptually distinct, policy questions:

1. How should the government fund a stochastic sequence of required government expenditure, such as wars?

2. How might the tax system be used to optimally share the risk of productivity shocks across generations?

3. To what extent should a "Social Security" system be constructed to redistribute from later, presumably wealthier, generations to earlier, presumably poorer generations, and how should that system respond to productivity shocks?

As discussed in the analytical section, we will see broadly that the government institutes more strongly redistributive policies when the IES—a proxy for the government's ability to borrow without altering interest rates too much—and the degree of inequality-aversion in the social welfare function are high; under these conditions, deviating from a balanced budget is the least expensive in efficiency terms and the most rewarding in equity terms. We will also see that each of these policy questions highlights different aspects of how the planner uses the instruments available to achieve his optimum.

For all applications, the uncertainty should be envisioned as worldwide. I have assumed a closed economy, which means that the smaller the country and the more local the shock, the better the insurance terms on worldwide capital markets. Shocks that are perfectly insurable abroad are not the focus of this analysis.

Calibration. Before engaging with these specific applications, I briefly describe the functional forms, general calibrations, and numerical methods I employ.[22]

In my benchmark simulations, I use household utility functions from Chari, Christiano and Kehoe (1994) and later Farhi (2010):

$$u(c, \ell) = (1 - \gamma)\log(c) + \gamma \log(1 - \ell),$$

where $\gamma = 0.75$. The pure rate of time preference for both households and the planner is 2% per year, or, since a period corresponds to roughly thirty years, $\Delta = \beta = 0.98^{30}$. In sensitivity analysis I will consider alternative values for the IES, which is not possible within the context of a log-log utility function. In such analyses, I use the following utility function:

$$u(c, \ell) = \frac{c^{1-\sigma}}{1-\sigma} - \eta\frac{\ell^{1+\varphi}}{1+\varphi},$$

where $1/\sigma$ varies and captures the IES, $1/\varphi$ represents the Frisch elasticity of labor supply which I set to 1/2, and $\eta = 1$.[23] I refer to the prior calibration as "log-log" utility, and the latter as "isoelastic" utility.

I use different functional forms for the social welfare function depending on whether the IES is 1. In either case, ζ parameterizes the degree of inequality-aversion (concavity) of the social welfare function, with $\zeta = 0$ representing a utilitarian planner and $|\zeta| \to \infty$ representing a "Rawlsian" planner who cares only about the worst-off generation. If

[22]The model developed in the previous sections and applied here is clearly highly stylized. Thus, these calibrations were not chosen to be especially "realistic," as such realism is beyond the scope of this paper and deserves its own attention in future work. Results presented in this section, therefore, should not be taken too literally as quantitative recommendations for actual policy, but rather as illustrations of the properties of optimal policy that are induced by the introduction of OLG to the Ramsey model.

[23]I alter the function to make it consistent with balanced growth by multiplying the disutility of labor by $\exp(at)^{1-\sigma}$, where a is the trend growth rate of labor-augmenting productivity.

IES = 1, I use the CARA function

$$W(u) = \begin{cases} u & \text{if} \quad \zeta = 0 \\ -\frac{1}{\zeta}\exp(-\zeta u) & \text{if} \quad \zeta > 0 \end{cases}$$

If IES $\neq 1$, I use the CRRA function, altered according to Kaplow (2003)

$$W(u) = \begin{cases} \frac{[(1-\sigma)u]^{-\zeta}u}{1-\zeta} & \text{if} \quad \zeta \neq 1 \\ \ln u & \text{if} \quad \zeta = 1 \end{cases}$$

To ensure the welfare function has its usual properties, ζ should be weakly positive if $\sigma < 1$, and ζ should be weakly negative if $\sigma > 1$. These choices make the welfare function scale invariant, which will play a crucial role in developing a recursive representation of the planner's problem when trend productivity growth is introduced.

Production is Cobb-Douglas with $\alpha = 1/3$ and depreciation of 8% per year, meaning

$$F(K, L; A) = K^{1/3}(AL)^{2/3} + 0.92^{30}K.$$

For simplicity and intuition, I consider only models that are purely dynastic or purely OLG. A "Ramsey" model sets $\delta = \beta, \underline{z} = -\infty$, and an "OLG" model sets $\delta = 0, \underline{z} = 0$. If markets are incomplete, I must specify the upper and lower debt limits; I follow Farhi (2010) and set them at 50% and -10%[24] of the average across states of the first best GDP for that state if it were absorbing. To avoid issues with the initial period, I always allow the system to run for 100 periods in the "good" state prior to any results that are shown. I will discuss the calibration of the stochastic process within the context of each application.

[24] Farhi (2010) chose 100% and -20%, but those values are for the amount of debt *issued*, and the interest rates are nearly double in this model due to the long period length.

Solution Method. To compute a numerical solution to the plannning problem, I first reformulate Problems 1.2.1 and 1.2.2 recursively, the details of which can be seen in Appendix A.4.1. I solve the resulting Bellman equations on rectangular, bounded state spaces, verifying that the bounds do not bind, and that altering them does not substantially alter the results.

My solution method is collocation using Chebyshev polynomials on a sparse grid.[25] Broadly, collocation techniques parameterize the value function as a weighted sum of N smooth functions (in this case, Chebyshev polynomials). Then, N points—the collocation points, or nodes—are chosen inside the state space. Finally, I solve for weights, using a value function iteration approach, such that the Bellman equation holds exactly at those N points if the value function is represented by the weighted sum. I choose the N points and the set of N Chebyshev polynomials according to the sparse grid method of Smolyak (1963).[26] Finally I ensure that the grid is sufficiently dense such that the Bellman equation holds quite closely at points other than the collocation nodes. Further details can be found in Appendix A.4.2.

1.5.1 Government Expenditure Shocks

Determining optimal financing of government expenditure shocks, such as wars,[27] has been a classic application of Ramsey optimal taxation models. In such standard models, shocks to government expenditure should be financed entirely through insurance if

[25]Excellent references include Miranda and Fackler (1997), Mark (2004). The classic book on numerical methods, including this one, is Judd (1998).

[26]Excellent references include Krueger and Kubler (2004) and Judd et al. (2014).

[27] There are multiple possible ways to model a war or other government expenditure in the context of the present model. For simplicity, I consider government expenditure to be a pure, required resource cost for the economy, having no effect on utility. On the other hand, one might consider government expenditure to be purchases of public goods, which surely enter household utility, but also are presumably optimally chosen rather than required and thus require a far richer model.

markets are complete (Chari, Christiano and Kehoe, 1994), and entirely through debt if markets are incomplete (Farhi, 2010). These stark benchmark policy prescriptions make this question an excellent laboratory for exploring the tax smoothing properties of the present OLG model.

Unanticipated Spending Shock

To build intuition, I first consider the simplest version of government spending uncertainty—a completely unanticipated spending shock (a "war"), known to last a single period, with zero government spending before or after. I assume that the economy is in steady state prior to the war[28] The war costs 10% of steady state GDP.[29]

Figure 1.3 shows the optimal policy response to this unanticipated shock for a constant IES of 5 for a variety of social planners.[30], while Figure 1.4 shows the optimal response for a utilitarian planner for a variety of IESs. Together, these figures illustrate the difference in optimal policy response between a Ramsey model and an OLG model, while highlighting through comparative statics the two economic forces introduced by OLG.

The clearest distinction is between the OLG models and the Ramsey model. In the

---

[28]For Ramsey models, which have a continuum of steady states, I choose the one with the same government debt as the OLG steady state to which it is compared.

[29] Given the completely unanticipated nature of the shock, and completely deterministic economy thereafter, there is no distinction between complete and incomplete markets, or between household holdings of capital and risk free bonds. Nonetheless, a choice must be made about how household balance sheets respond to the unanticipated shock in the period of the shock. I make the simplest possible assumption, which is that households officially hold risk free debt (paying $1/\beta$ gross interest) in the pre-shock steady state; thus, old consumption in the period of the shock is bounded below by its steady state level in the OLG model. This eliminates the ability of the planner to decrease old consumption by decreasing young labor supply, which makes the solution more realistic.

[30]I have chosen this (perhaps unrealistically high) IES to allow the government to raise an unlimited amount of revenue from debt issuance and thereby give the government a "real choice" over how to finance the war in the short run. Similar properties could be generated by allowing the government to borrow from abroad, or choosing a more elastic production function.

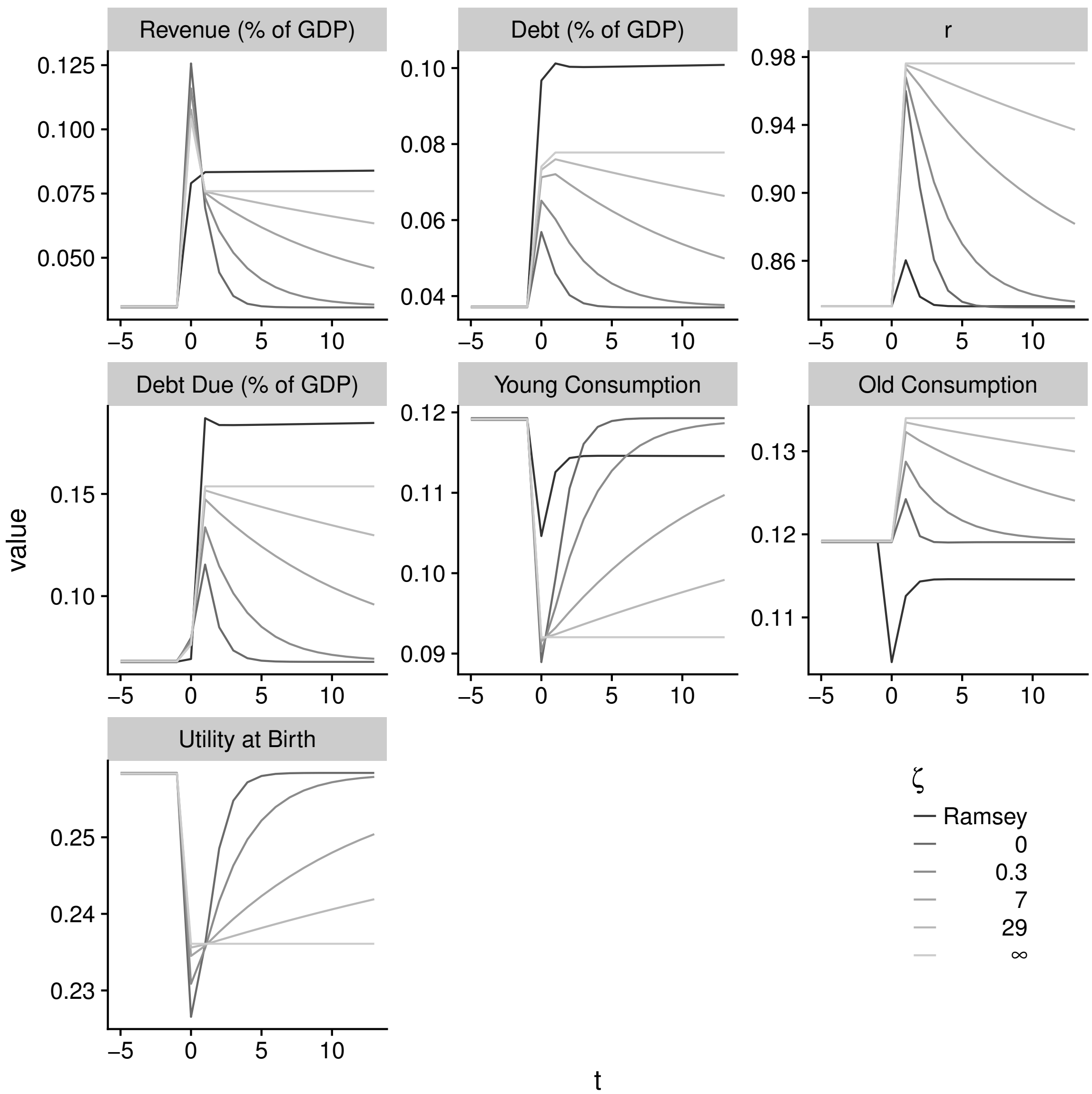


Figure 1.3: Optimal policy response to a completely unanticipated government spending shock equal in size to 10% of steady state GDP. These graphs show many economic variables each period, where period 0 is the period of the shock. Multiple models are considered: a Ramsey model, along with several OLG models featuring different planners with different levels of inequality-aversion, parameterized by ζ. All models assume isoelastic utility with an IES of 5.

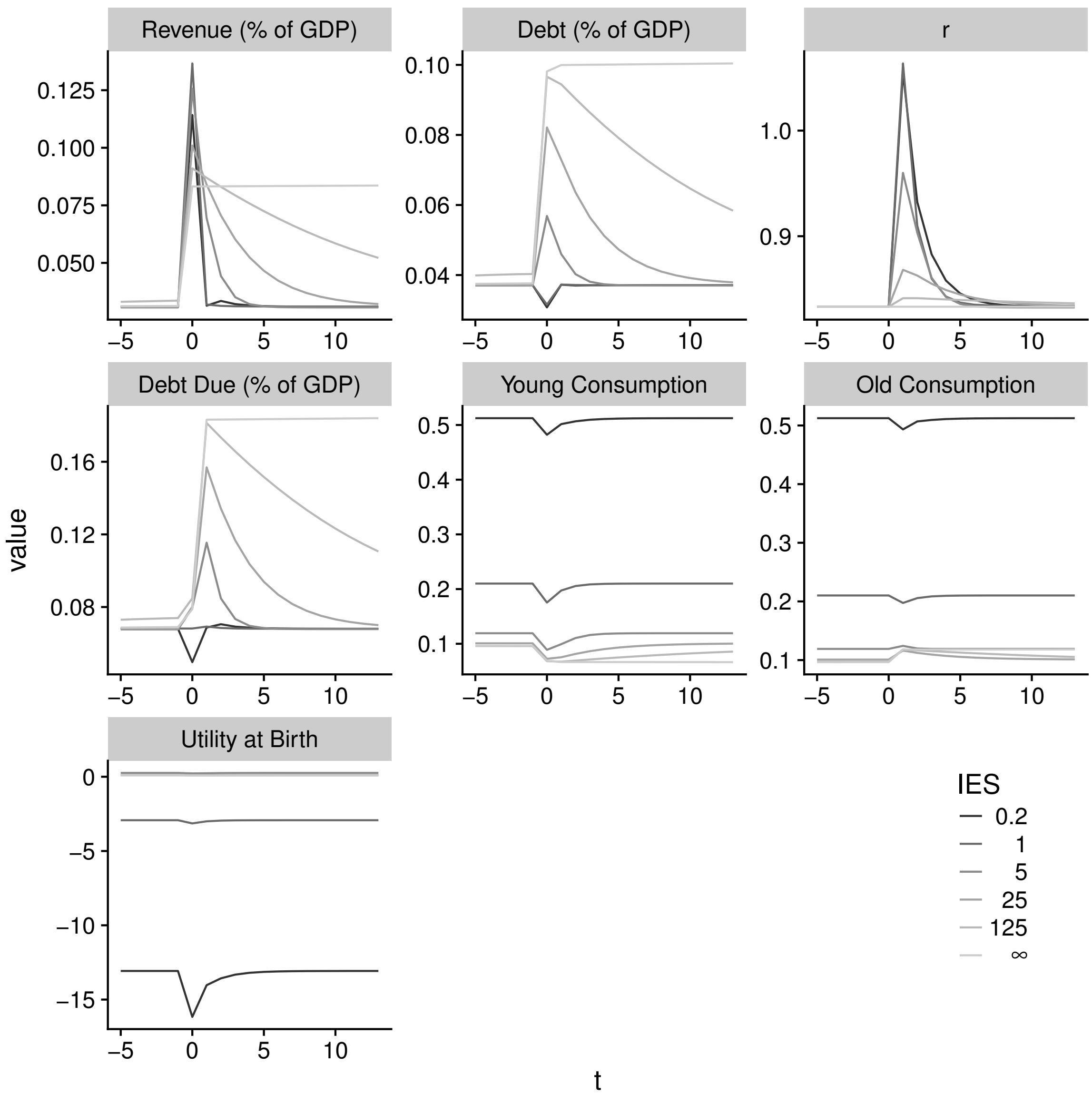


Figure 1.4: Optimal policy response to a completely unanticipated government spending shock equal in size to 10% of steady state GDP. These graphs show many economic variables each period, where period 0 is the period of the shock. The figure compares OLG models featuring a utilitarian planner but a variety of IESs.

Ramsey model, the war is paid for over the rest of history, with taxes and debt constant from period 2 onward at a much higher level than previous to the war. On the other hand, in all OLG models except the ones featuring a Rawlsian planner or an infinite IES, taxes are higher during the war than afterward, with taxes eventually returning to the pre-war level.

The reason for the temporary increase in taxation is the loss of Ricardian equivalence, through two different channels. First, in an OLG model, financing the war with debt drives up the interest rate, which increases the amount of revenue that must be raised later and the associated efficiency cost. This effect is strongest when the IES is low, since interest rates must move substantially to induce households to hold more government debt.[31] Second, in an OLG model, only contemporaneous taxation increases labor supply through the income effect, which makes taxing during the war, when resources are scarce. This effect is strongest when the income effect is strongest, which in this calibration is also when the IES is low. The price of this more efficient, contemporaneous taxation is inequality across generations, which is visibly falling with the IES.

On the other hand, the planner's desire for equality across generations pushes back toward more constant labor tax rates. In this model, in which productivity is constant, labor taxes play a large role in determining the welfare of a given generation. Thus, a more inequality-averse planner cares more about maintaining a constant tax, which we see in Figure 1.3.[32] Additionally, a more inequality-averse planner reverts more slowly to steady-state taxation because he attempts to compensate generations that

[31] This should not be interpreted strictly as a statement about the IES; rather, it shows that the government will finance the war with more debt if it has freer access to debt markets. This could occur through easier crowding out of consumption (higher IES), investment (greater elasticity of substitution between labor and capital), or net exports (greater ability to borrow on worldwide capital markets in an open economy generalization).

[32] The reason for the spike in tax rates at date 0 even for a Rawlsian planner is that capital also affects welfare by raising the pre-tax wage. An inequality-averse planner thus places higher taxes on generations with access to more capital—in this case, the generation born at date 0.

faced high taxes in youth with more consumption in retirement; to do this, he offers a better interest rate on debt, which in turn requires raising taxes on the subsequent generation. Inequality across generations is, intuitively, falling with the planner's degree of inequality aversion.

Stochastic Government Spending

Next, I simulate a complete markets model with more conventionally stochastic government spending, which allows some preparation for shocks. I assume i.i.d., equal probability draws of government spending that is 20% ("peace") or 25% ("war") of steady-state GDP and log-log utility. Simulations with other IESs can be found in the Appendix; they are qualitatively similar and demonstrate the same comparative statics as with unanticipated spending shocks above. I simulate three cases: an OLG model with a utilitarian planner, an OLG model with a concave planner ($\zeta = 24$), and a Ramsey model as a benchmark.

I begin by presenting graphs (Figure 1.5) showing the evolution of several variables during state transitions. This is done by alternating between four periods in the low spending state and four periods in the high spending state.[33]

I begin with the Ramsey model. We see the familiar result that shocks to the budget are covered almost entirely using insurance (decreased payouts on the debt), leaving labor taxes and debt to GDP ratios almost constant.

In the OLG model with a utilitarian planner, the most important change is that insurance is much less used, for both efficiency and equity reasons. From an efficiency point of view, insurance in a Ramsey model is desirable. When the governemnt takes advantage of it in the event of a war, it acts as a negative wealth shock to households, which causes them to efficiently work harder through the income effect. But in an

[33] Of course, this pattern is not anticipated by the planner.

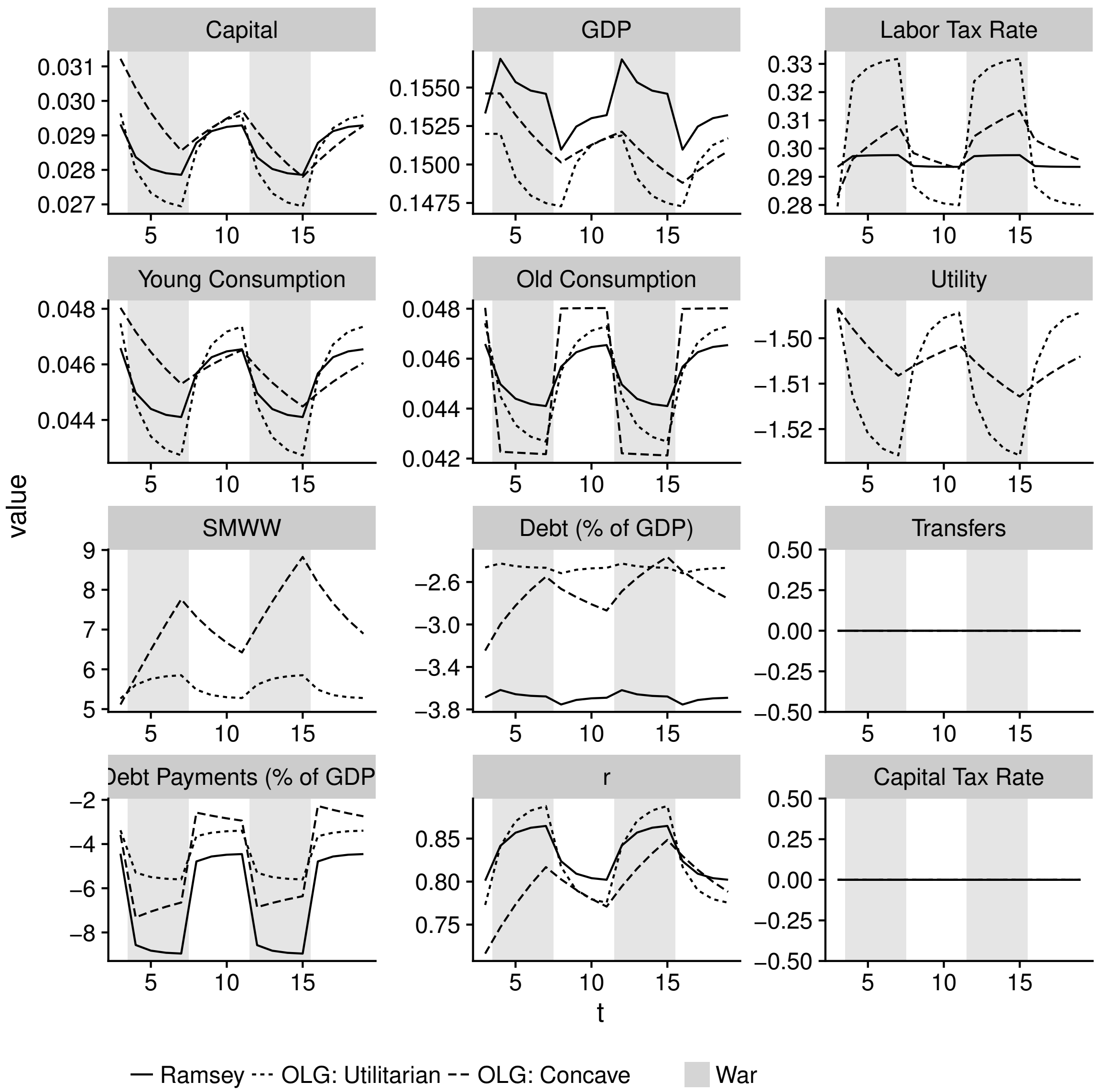


Figure 1.5: This figure captures the salient features of state transitions in the complete markets model of stochastic government expenditure when utility is log-log. Shading marks periods of high government spending ("wars"). The concave planner has $\zeta = 24$. The model was first allowed to run for 100 periods in state 1 to arrive at a steady state. The Ramsey model was started with debt equal to -2.4% of nonstochastic steady state GDP, which corresponds to the level of government debt in the OLG model after 100 periods of low government spending; initial conditions are irrelevant for the OLG models.

OLG model, the negative wealth shock accrues to households that have already retired, eliminating this effect. From an equity point of view, insurance places risk entirely on the generation that is retired at the time a war arrives, meaning the government will not want to fully insure the war. Instead of insurance, the government uses large variations in the labor tax rate to cover shocks to the budget.

Moving to the OLG model with a concave planner, we see the planner making use of all fiscal instruments available to distribute the risk of a war across all generations. Increased use of insurance places risk on retirees; increased labor taxes place risk on workers; borrowing during the war places risk on unborn generations; and the accumulation of a better fiscal position up front places risk on generations that are already deceased.

Figure 1.6 shows the results of randomly simulating the model across ten thousand periods, and offers a different view of the tax smoothing properties of these models. The Ramsey model features tax rates that are quite close to constant. On the other hand, the OLG models feature labor tax rates that are decidedly non-constant, but significantly autocorrelated conditional on s_t and s_{t-1}. The fact that they are non-constant is attributable to the loss of Ricardian equivalence, but the autocovariance is a form of tax smoothing; since it appears particularly strong in the concave case, we may conclude that this tax smoothing exists primarily for equity reasons rather than the usual efficiency reasons.

1.5.2 Productivity Shocks

Next, consider an economy in which productivity is uncertain but has no trend growth, and in which there are no required government expenditures. If households are dynastic, the welfare theorems apply, and there is no need for any intervention by the government. However, in an OLG model, the planner faces a problem of optimally sharing the

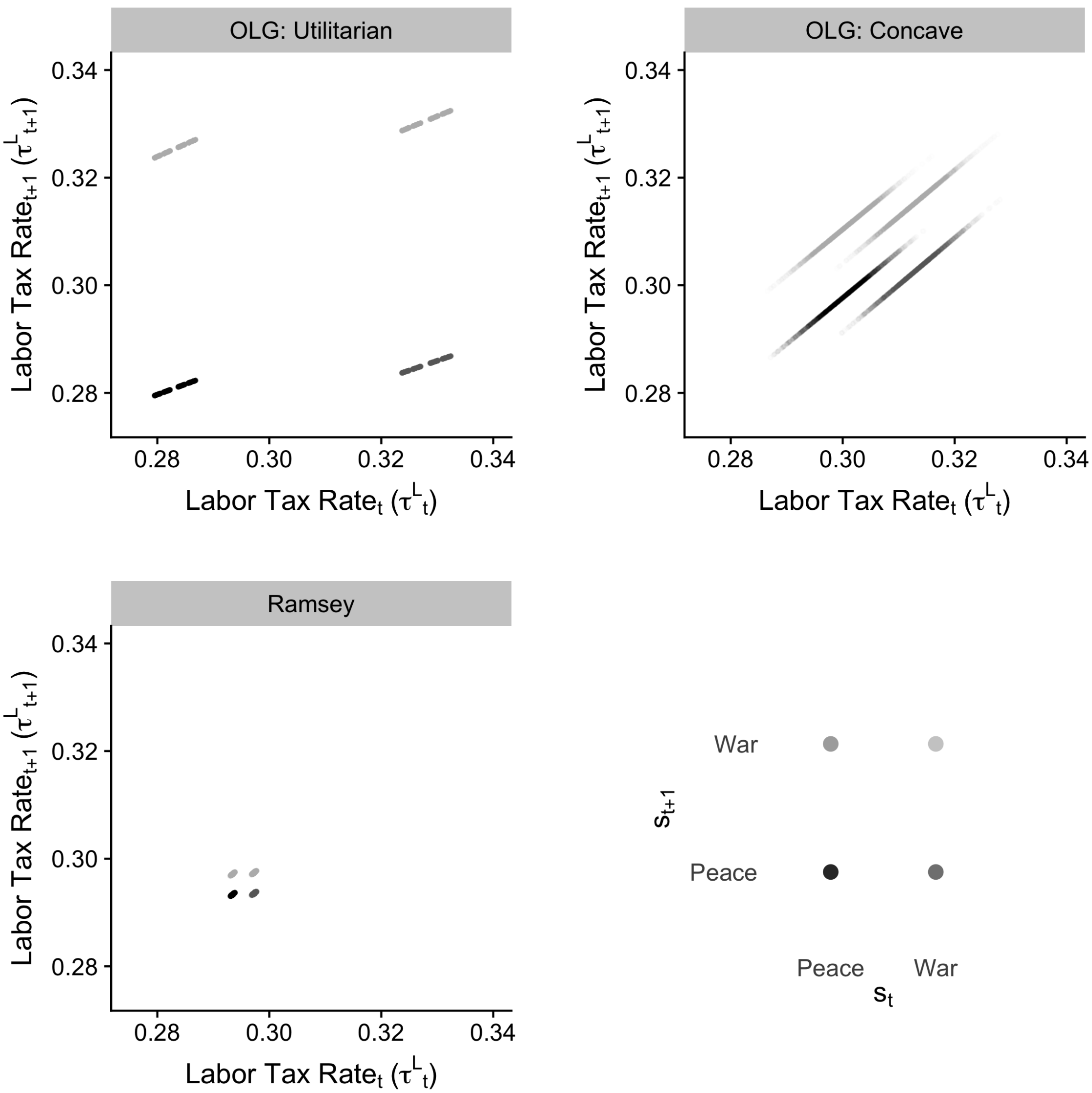


Figure 1.6: This figure presents the results of a random simulation of the complete markets model of stochastic government expenditure when utility is log-log. The model is simulated over 10000 periods (after a 100-period startup period that is excluded), and each point represents one period. The concave planner has $\zeta = 24$. The labor tax rate at date t is on the horizontal axis, and the labor tax rate at date $t+1$ is on the vertical axis. The different shades of gray represent different sequences of shocks received at date t (s_t) and date $t+1$ (s_{t+1}). The Ramsey model was started with debt equal to -2.4% of nonstochastic steady state GDP; initial conditions are irrelevant for the OLG models.

productivity risk across generations. Since the government has no other purpose besides this intergenerational redistribution, we can interpret all government payouts on debt as Social Security (SS) payouts.

Therefore, the existence of complete markets is isomorphic to the existence of defined contribution Social Security—a retirement program in which payouts may vary with the state of the world, but in a manner that is proportional to the amount contributed or saved. If markets are incomplete, then the government must make the same payment on its debt in every state of the world, and thus can only vary its payouts to retirees through the use of lump sum transfers, which we will see are inefficient.

The producitivity shocks are calibrated as equal probability i.i.d. draws of $A = 1.08$ and $A = 0.92$[34]; I increase the shocks to $A = 1.12$ and $A = 0.88$ for the incomplete markets example to make the lump-sum transfer instrument active.

Incomplete Markets

I begin with the incomplete markets model, in which the government lacks access to a defined contribution SS system, which might be more intuitive to readers in the U.S. Figure 1.7 shows the evolution of various variables during state transitions, and captures the essence of the government's policies. The first property to notice is the very minor use of lump sum SS. This is because such sump sum transfers are highly inefficient; they are granted regardless of labor supply in the previous period and thus disincentivize labor via the income effect. Since the utilitarian planner wishes to avoid distorting capital accumulation to redistribute across periods, he must make use of countercyclical labor taxes to compensate for a decrease in the tax base during the downturn. The concave planner cares more deeply about providing insurance to the

[34] 16% is approximately the worst underperformance of trend over a 30 year period during the previous century or so.

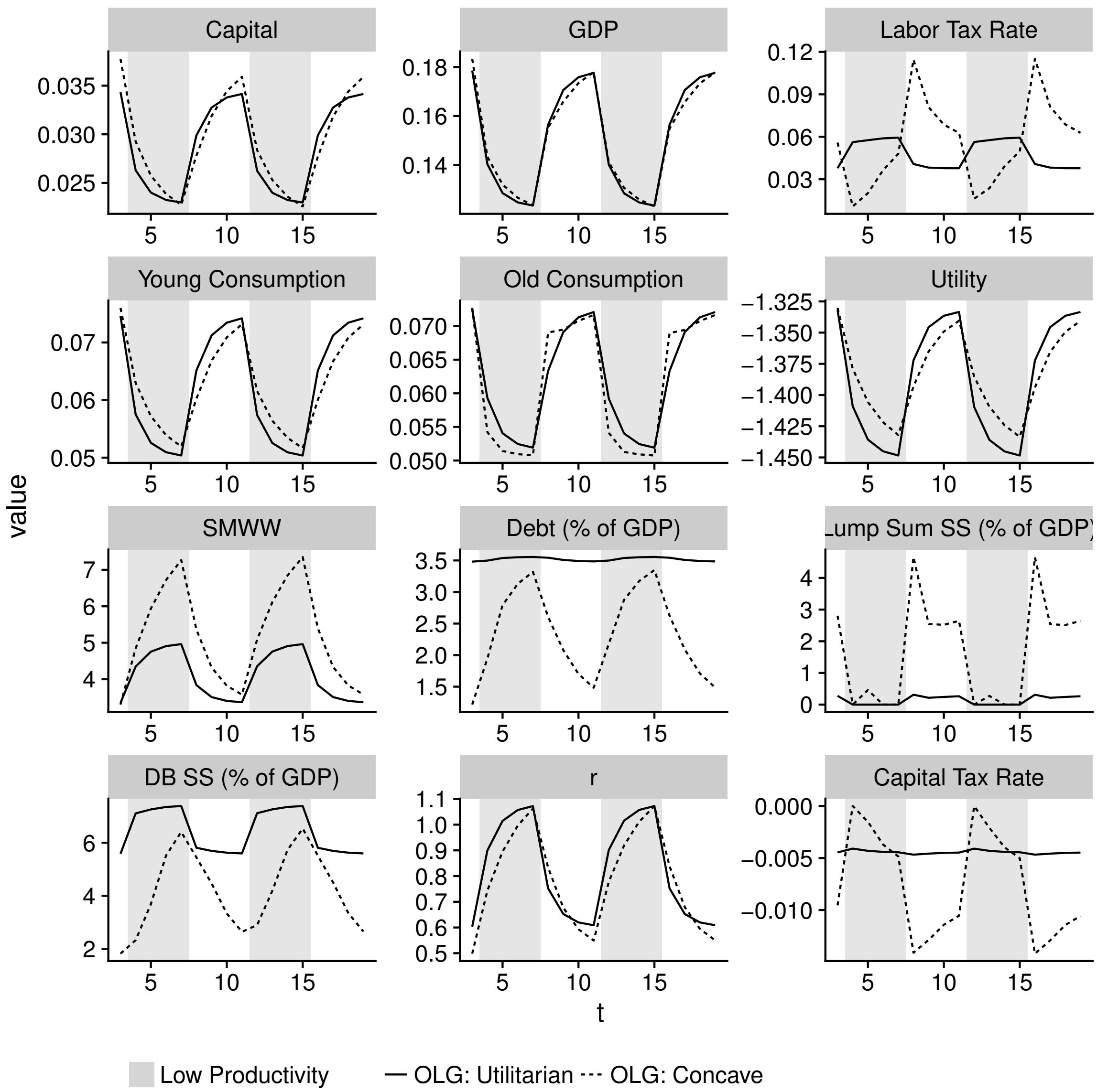


Figure 1.7: This figure captures the salient features of state transitions in the incomplete markets model of stochastic productivity when utility is log-log. Shading marks periods of low productivity. The concave planner has $\zeta = 4$. The model was first allowed to run for 100 periods in the good state to arrive at a steady state.

unborn, and so uses strongly procyclical labor taxes; the lost revenue is replaced with steeply increasing borrowing during downturns, since varying payouts to retirees would require using the inefficient lump sum SS instrument during good periods.

Complete Markets

Now suppose the government has access to a defined contribution Social Security system. Figure 1.8 shows the evolution of various variables during state transitions, and captures the essence of the government's policies. If the planner is utilitarian, then the desire for intergenerational is primarily static and focused on evenly distributing the resources available in a given period. Since retirees already share in productivity shocks to the extent that capital returns are affected, policies such as tax rates, SS payouts as a fraction of GDP, and debt to GDP ratios are nearly constant; most variables move in concert. As in the incomplete markets case, an explicitly inequality averse planner wishes to compensate households born into poor states of the world with a lower tax rate.[35] In this case, such a reduction in revenue is funded through a sharp reduction in SS payments.

Perhaps counterintuitively, SS payments do not insure the old against shocks to their retirement income—a common justification for SS in policy debates—but instead do precisely the opposite. However, recall that the social planner's welfare function was constructed to respect private preferences over risk. Thus, the planner has no motivation to decrease the amount of risk assumed by retirees. Rather, it is socially optimal for retirees to assume *more* risk than is privately optimal, so as to reduce the risk borne by future generations. This is especially true in the first period of a low

[35] and, during the first period of a low productivity spell, very small lump-sum transfers to young households

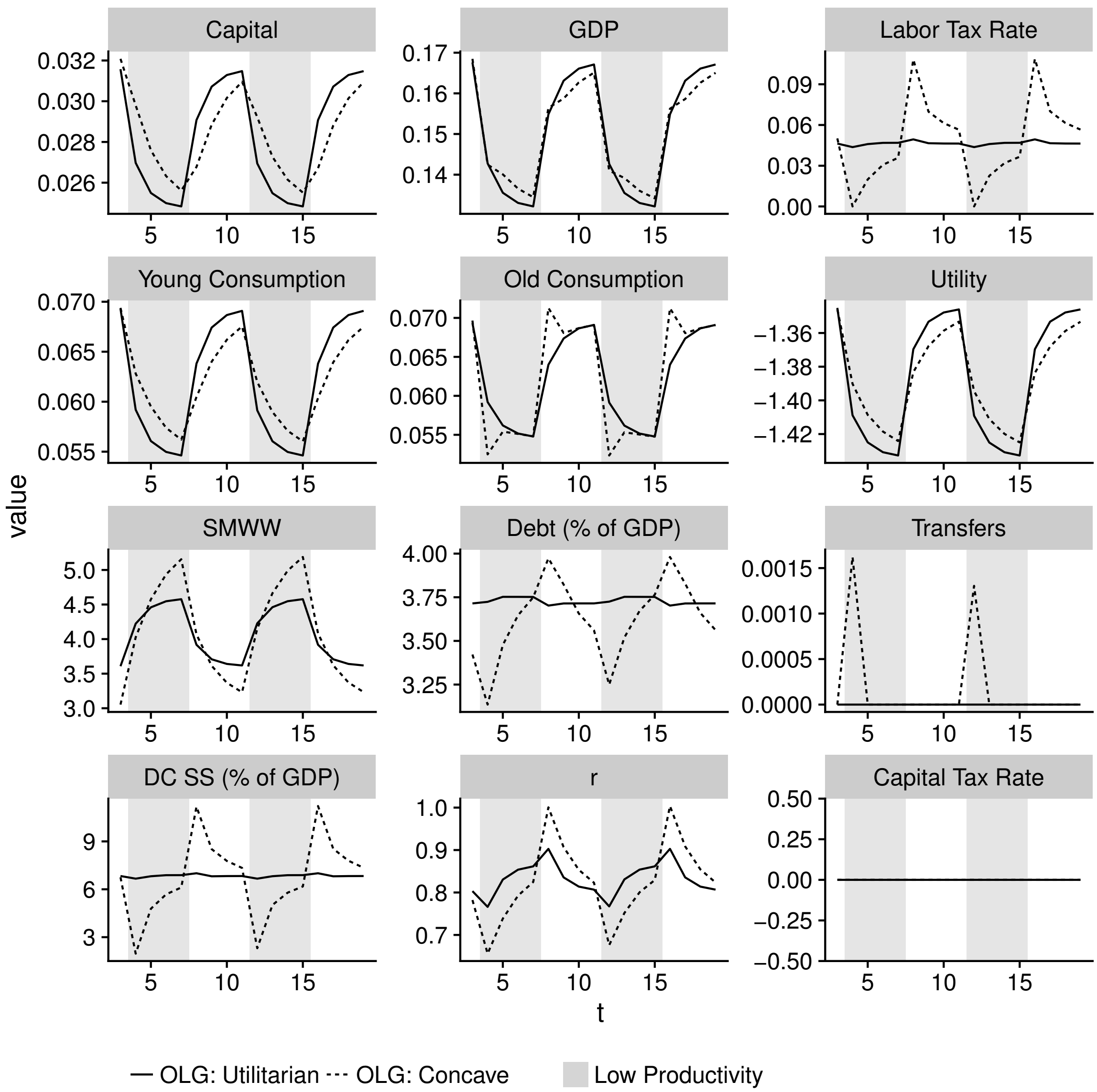


Figure 1.8: This figure captures the salient features of state transitions in the complete markets model of stochastic productivity when utility is log-log. Shading marks periods of low productivity. The concave planner has $\zeta = 4$. The model was first allowed to run for 100 periods in the good state to arrive at a steady state.

productivity spell[36] when the planner is concave, since retirees in such a period will have enjoyed high utility during youth, thereby reducing their social marginal welfare weight.

1.5.3 Productivity Growth

A common argument for an intergenerational transfer system like Social Security in the U.S. is that, if productivity is growing on average, then future generations will be wealthier than current ones. Thus, a system that redistributes from future generations to earlier ones is socially desirable. By adding a deterministic trend to productivity, the present model is well-suited to analyzing how such a system might be optimally designed. Preserving shocks to productivity around such a trend from the previous application allows analysis of how such a SS system should respond to productivity shocks—a major policy question at hand today, after a period of slower than average productivity growth. Toward this end, I add deterministic growth in A of 2.1% per year (the long run U.S. average) to the previous model.

Figure 1.9 depicts optimal policy in this model if the government has access to state-contingent debt. While preserving all properties of policy in the previous model, trend growth introduces several new effects. First and most importantly, labor taxes and SS payments are much higher, especially when the planner is inequality-averse. This is evidence of precisely the intuition stated above—redistribution from richer, future (younger) generations to poorer, current (older) generations, the motive for which grows with the planner's inequality aversion. This comes at a price however—a higher interest rate, which in turn crowds out capital formation and lowers utility, all the more when the planner is more inequality averse. It is worth noting that the intergenerational

[36] and vice-versa

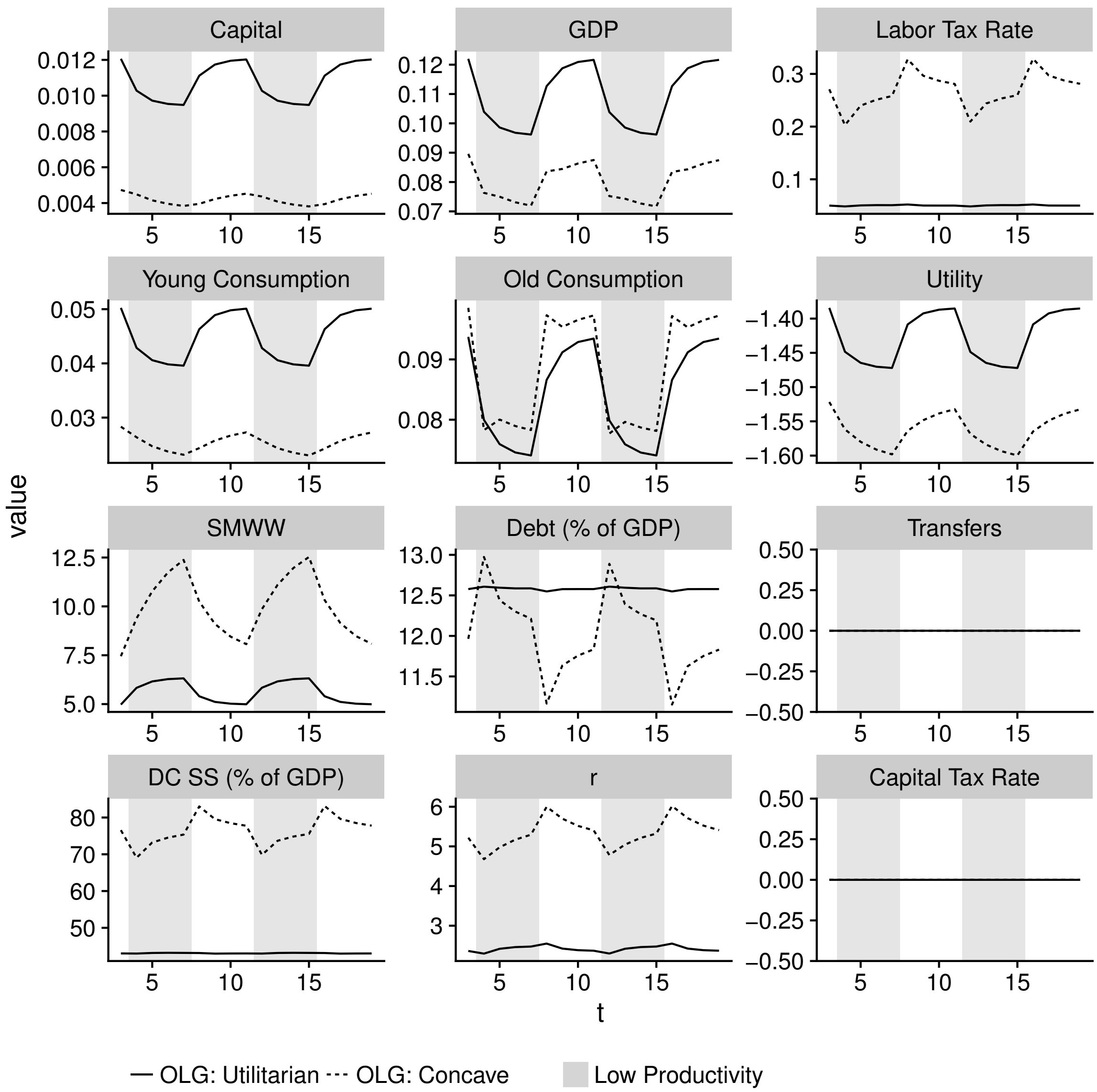


Figure 1.9: This figure captures the salient features of state transitions in the complete markets model of stochastic productivity with trend growth of 2.1% per year when utility is log-log. Shading marks periods of low productivity. The concave planner has $\zeta = 4$. The model was first allowed to run for 100 periods in the good state to arrive at a steady state. All variables detrended, with old consumption and debt (SS) payments detrended by the previous period's expected productivity.

risk-sharing carried out by state-contingent SS payments is amplified by the trend growth. This is because, with this calibration, the "stakes" are higher—fluctuations are 8% above or below trend, which are very large relative to the earnings of the previous generation.

Figure 1.10 depicts optimal policy in this model if the government does not have access to state-contingent debt. Importantly, trend growth does *not* preserve all properties of policy in the previous model in this incomplete markets case. Specifically, the intergenerational risk sharing that was implemented via lump-sum transfers has disappeared. When labor tax rates are higher, which they are here to redistribute the trend growth in productivity, the income effect of lump-sum transfers has a larger effect on the government budget and, therefore, a larger efficiency cost. Thus, the planner decides to forgo intergenerational risk sharing. This highlights an important tradeoff between intergenerational risk sharing and deterministic intergenerational redistribution faced by a planner without access to state-contingent debt, but not by a planner with access to state-contingent debt.

1.5.4 Summary

Though each of these applications gives rise to substantially different optimal policies, the ideas are similar. An OLG planner, like a Ramsey planner, cares about smoothing distortionary costs of taxation, but also cares about intergenerational equity to a degree parameterized by his inequality-aversion. Income effects, however, often impair his ability to achieve such equity, and are higher for lower IESs with constant-IES utility functions.

One final note: Access to state-contingent debt changes policy substantially—especially for the applications to productivity shocks and growth—which prompts a discussion of whether such a policy instrument is realistic. Traditionally, in Ramsey

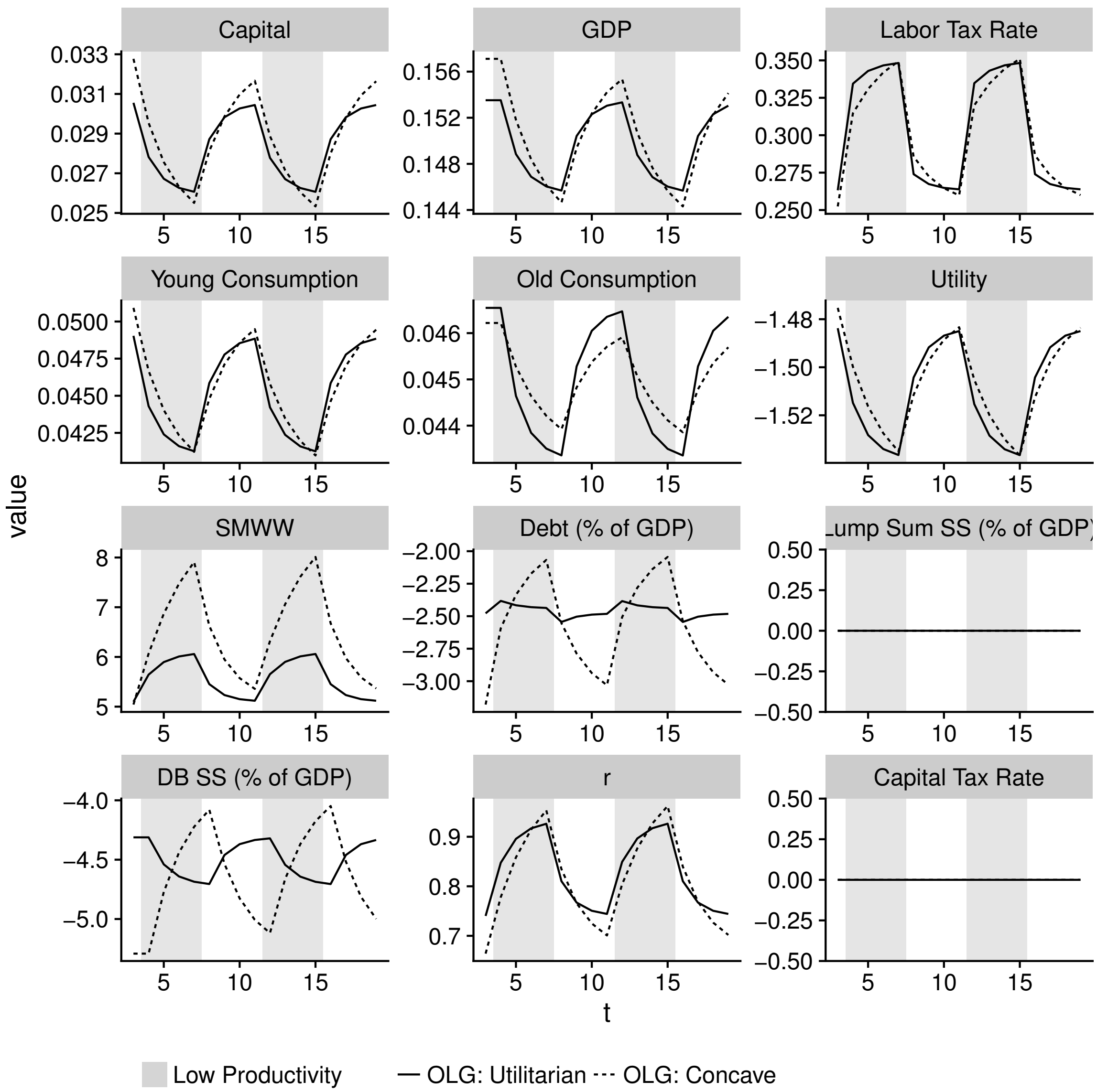


Figure 1.10: This figure captures the salient features of state transitions in the incomplete markets model of stochastic productivity with trend growth of 2.1% per year when utility is log-log. Shading marks periods of low productivity. The concave planner has $\zeta = 4$. The model was first allowed to run for 100 periods in the good state to arrive at a steady state. All variables detrended, with old consumption and debt (SS) payments detrended by the previous period's expected productivity.

models focused on government spending shocks, such state-contingent debt is widely seen as an unrealistic ability for the government to default on its debt in bad states of the world in a manner that is anticipated and not penalized via future difficulties obtaining credit; it is assumed for mathematical convenience. However, I argue that in the context of productivity shock and growth applications, such state-contingent debt is quite realistic; it represents the government's ability to issue different-sized SS checks depending on the state of the world. Though U.S. law does not explicitly specify SS as a state-contingent benefit that varies with the economoy's performance in late middle age or after retirement, it is quite reasonable to think that those benefits will be revised upward via new legislation if the economy performs unexpectedly well. The reverse seems unlikely due to loss aversion and political economy concerns, which are beyond the scope of this paper. But overall, the ability of the government to change SS payments in a proportional manner seems far more realistic than issuing true lump-sum transfers to retirees if the economy performs unexpectedly well.

1.6 Robustness: More than Two Generations

The two-generation model considered to this point has limited concrete applicability, as each period represents roughly thirty years. To check whether the foregoing results are qualitatively robust to a greater number of generations, and therefore shorter period length, I extend the complete markets model to feature a generic number of generations, but no bequest motive for simplicity. As this is merely a robustness test, I focus only on the first application: a model with stochastic government expenditure, and constant productivity. I consider only the simple, complete markets case with a utilitarian planner.

Most aspects of the model are unchanged. Instead of two generations, one working and one retired, there are now J_2 generations, of which J_1 are working and $J_2 - J_1$ are retired, indexed by $j = 1, 2, ..., J_2$. Households' preferences and budget sets are precisely as before, properly extended to account for the further generations. The planner still has access to linear taxes on labor, state-contingent linear taxes on capital, age-contingent lump-sum transfers (which can without loss of generality be assumed to only be given to the youngest generation), and state-contingent debt.

Though the models are similar, an important complexity is added: Multiple agents trade in the same markets (for labor and state-contingent securities) at the same time, and an allocation is only implementable if it is compatible with all trading agents facing the same prices. In light of this complexity, I follow Werning (2007) and use the "market weights" approach to characterize the implementable allocations and formulate the planner's problem. Policy equilibria, implementable allocations, and the planner's problem are formally defined and proven in Appendix A.6.

1.6.1 Calibration

For results to be comparable, care must be taken when altering the number of generations, and implicitly the length of a period, to leave fundamental economic parameters unaltered. With this in mind, I choose the same log-log calibration of utility and Cobb-Douglas calibration of production, while changing the stochastic process so that the following are kept constant:

- The probability of a war per unit time
- Expected government expenditure as a percentage of steady-state GDP per unit time
- The variance of government expenditure as a percentage of steady-state GDP per unit time

Intuitively, this means that wars become less likely but more costly as the length of a period is shortened. β, Δ, and depreciation are also revised accordingly. I consider models with 3 generations (2 working), 4 generations (2 working), and 5 generations (3 working).

1.6.2 Results

Figure 1.11 shows how various variables evolve over state transitions in models with two through five generations. The results are very qualitatively similar to each other, suggesting that the nature of the system does not qualitatively depend on the number of generations. This provides reassurance that the assumption of two generations is not qualitatively important.

There are two qualitative changes. First, government debt is not always negative in steady state, which calls into question the "buffer stock" intuition of a risk-averse planner. However, the planner is also interested in avoiding dynamic inefficiency, which may require positive steady state debt depending on the discount rate and return to capital, which vary across these models. Second, capital taxes are no longer always zero, consistent with the intuition of Erosa and Gervais (2002). In an OLG model, the planner will generically wish to tax different generations at different rates; if that is forbidden, as here, capital taxes may fill a similar desire to redistribute across generations.[37]

Figure 1.12 shows the autocorrelation structure of labor taxes in models with more generations. Models involving 2, 3, or 4 generations have similar results, though shorter periods lead to slightly lower variation in optimal tax rates. Additionally, a greater number of generations and a smaller discount rate means that a longer history is

[37] Erosa and Gervais (2002) find that capital taxes are indeed zero if preferences are homothetic over consumption at different ages and separable from labor, as the log-log utility function is. However, their model features no uncertainty, and so ignores transitory desires to redistribute between generations stemming from shocks hitting some generations harder than others.

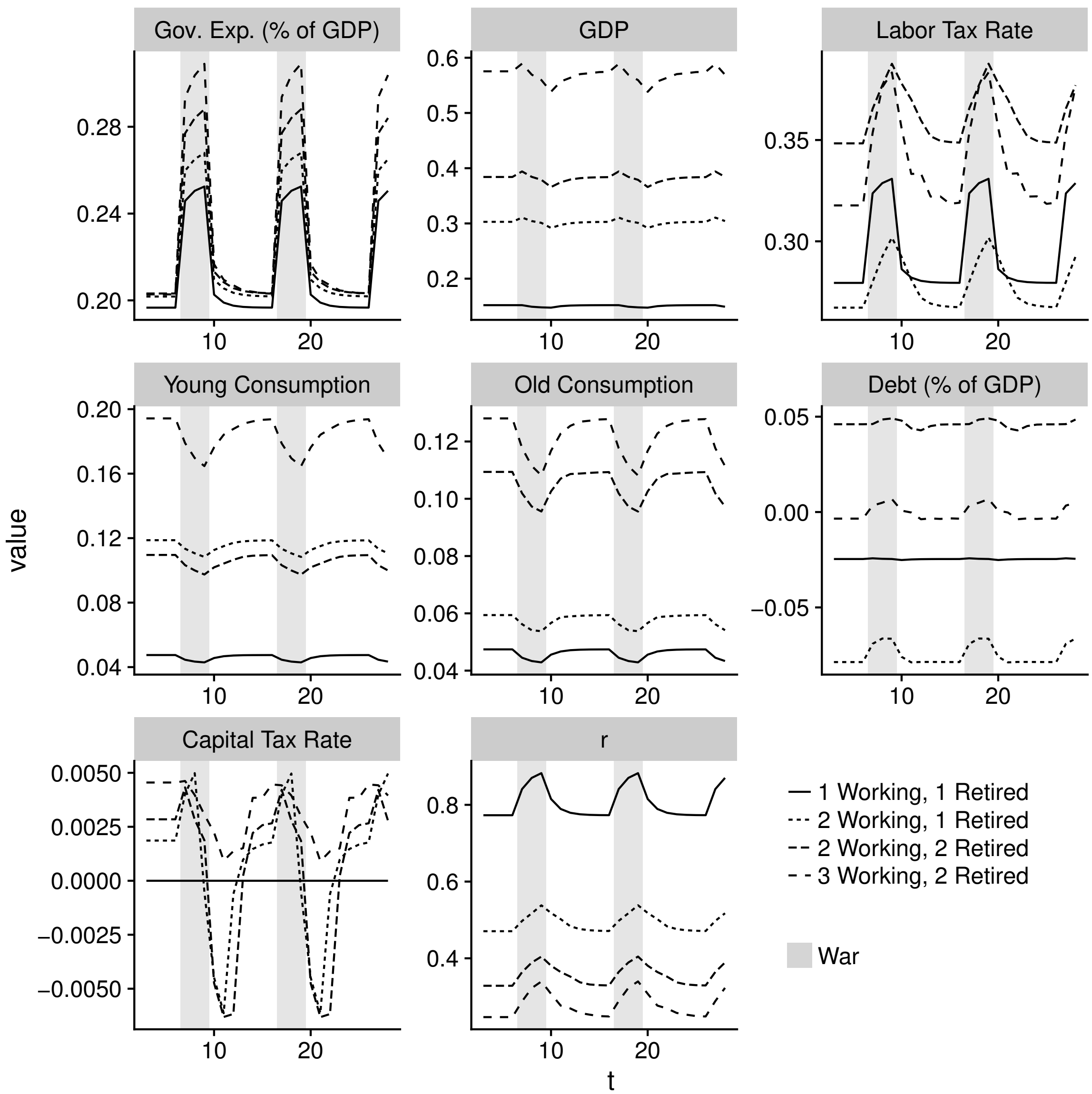


Figure 1.11: This figure captures the salient features of state transitions in the complete markets model of stochastic government expenditure when utility is log-log and the planner is utilitarian, with the number of generations ranging from 2 to 5. Shading marks periods of high government spending (“wars”). The model was first allowed to run for 100 periods in state 1 to arrive at a steady state.

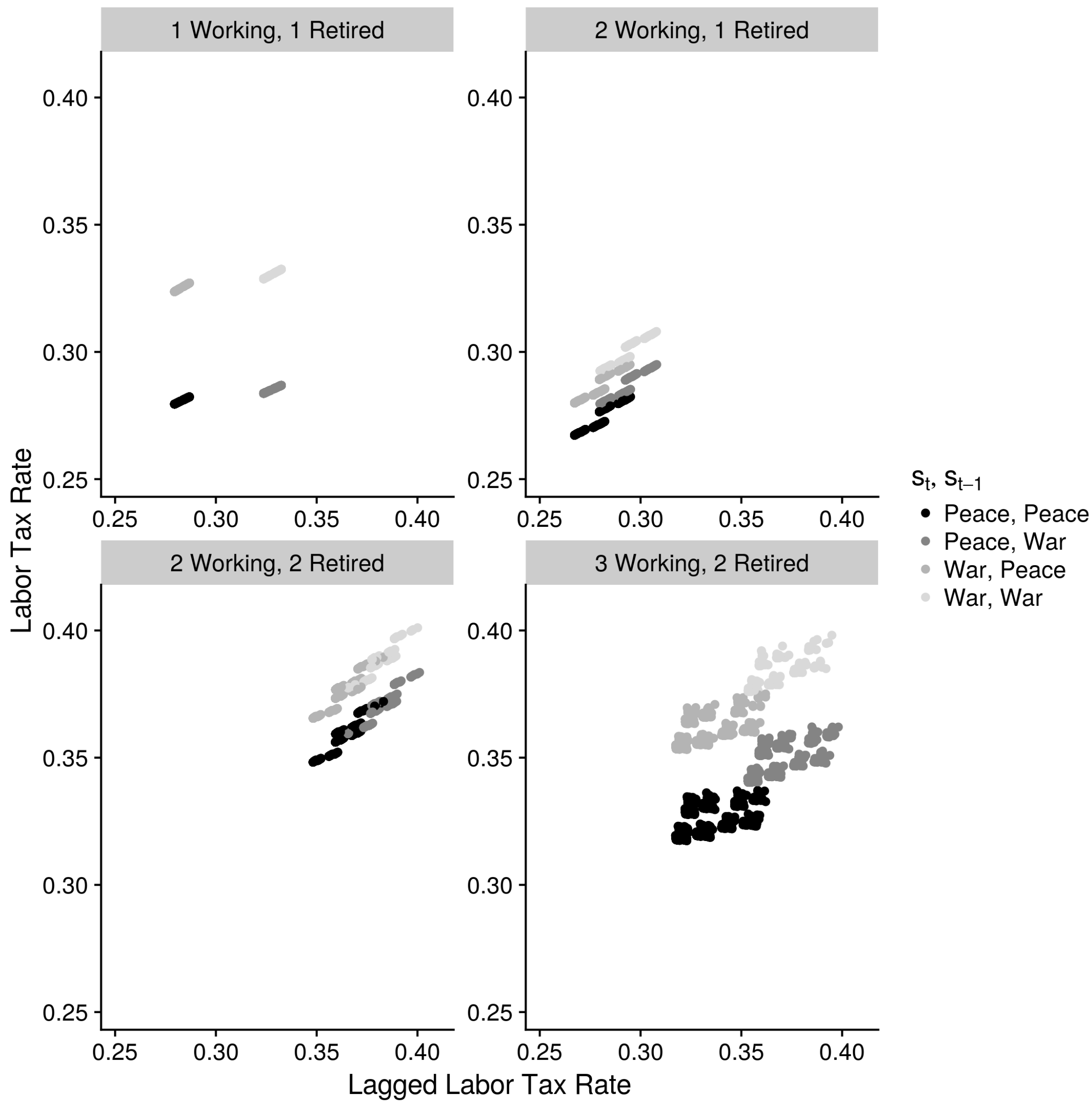


Figure 1.12: This figure presents the results of a random simulation of the complete markets model of stochastic government expenditure when utility is log-log and the planner is utilitarian, with the number of generations ranging from 2 to 5. he model is simulated over 10000 periods (after a 100-period startup period that is excluded), and each point represents one period. The lagged labor tax rate is on the horizontal axis, and the current labor tax rate is on the vertical axis. The different shades of gray represent different sequences of shocks received at date t (s_t) and date $t+1$ (s_{t+1}).

relevant. The model with 5 generations is, however, noticeably different, having more of a cloud-like structure rather than a tree-like structure. This noise could come from two different areas. First, this model, having six continuous state variables in its recursive representation, is subject to larger numerical error than the others. But more economically significantly, in this model, a single lag of the labor tax rate fails to capture all of the tax rates previously faced by currently working generations. This could explain the higher variation in the tax system.

Overall, the results of this section suggest that the assumption of two generations throughout most of the paper is not qualitatively important, and that the results may be applied to models with more generations and shorter periods.

1.7 Conclusion

This paper has developed a framework for analyzing optimal policy in a macroeconomic model with overlapping generations, aggregate shocks, and a planner who has available only linear taxes on labor and capital. This setup complicates a standard Ramsey taxation model by adding a redistributive motive across generations—and, therefore, a tradeoff between equity and efficiency that is ubiquitous in the public finance literature—as well as removing Ricardian equivalence. We saw analytically that optimal policy in this model requires that the government equate *total* costs of taxation over time, rather than purely focusing on *distortionary* costs.

These total costs fall into four distinct categories: First is the standard substitution effect—the distortionary cost. Second is a mechanical effect of income loss, which is constant over time in a standard dynastic economy with homogeneous agents, but cannot be assumed so in a model with redistributive motive. Third is the static income effect; higher taxes reduce income which in turn encourages labor supply and helps the government budget constraint. Fourth is a dynamic income effect which results from the effect of a proposed policy perturbation on asset prices and, therefore, perceived wealth. These two income effects can safely be ignored in the standard dynastic economy, since they are felt today even if they occur in the future; the same cannot be said of an OLG economy. Finally, these costs must be weighted by a social marginal welfare weight that captures redistributive preferences.

After the analytical investigation, I numerically applied the model to three distinct policy problems, which highlighted different aspects of these above findings. In the first, the government must decide how to optimally finance a stochastic sequence of required expenditure; optimal policy exhibits *partial* rather than perfect tax smoothing, with smoother taxes for higher IESs and more concave welfare functions. In the second, the government seeks to intergenerationally distribute the risk associated with a sequence of productivity shocks; optimal policy uses procyclical Social Security payments, if available, to ensure that retirees share in these productivity shocks. Finally, I added trend productivity growth, which creates a desire to systematically redistribute from later generations to earlier ones; this is achieved through a Social Security-like system that involves high government debt and high labor taxes—all the more so when the social welfare function is more concave. These numerical results qualitatively match how developed countries actually behave, suggesting that policy is on the right track.

Finally, we saw that these results are qualitatively robust to the addition of further generations and associated shortening of the period, at least for a benchmark case.

1.7.1 Future Work

This paper suggests numerous directions for future research. Perhaps most importantly, policy recommendations could become more numerically realistic in a model with a substantially shorter period and, as a result, substantially more generations and a richer distribution of possible shocks. Unfortunately, the difficulty of solving such a model normally grows exponentially in the number of generations, as each generation adds dimensions to the state space of the recursive representation; my sparse-grid method reduces this growth to cubic, though that still becomes quite rapid. A likely solution to this curse of dimensionality would be a switch to continuous time and an adoption of a framework akin to the perpetual youth model of Blanchard (1985). While this is not straightforward in a model with elastic labor supply, it seems more promising than the gradual addition of generations in a discrete framework. Alternatively, one might keep a discrete set of generations with different life expectancies, but solve the model approximately using distributions as state variables, similar to Krusell and Smith, Jr. (1998).

A second obvious extension of the work regards intra-cohort heterogeneity and, therefore, a motivation for redistribution within a single cohort. Such assumptions would allow dropping the linear taxation assumption and, instead, a switch to a Mirrleesian framework. Ideally, such a study would examine an economy facing both aggregate *and* idiosyncratic uncertainty and begin to address the issue that redistribution may be most desirable precisely when it is least affordable.

Third, one could consider other applications of the model. These might include fertility shocks, shocks to asset returns that do not affect labor productivity, or innovations to the growth rates of variables rather than transitory shocks to levels.

Fourth, the model presented in this paper, together with the characterization of optimal policy, raises empirical questions. While the response of taxable income to contemporaneous labor tax rates is a heavily studied topic,[38] the response of asset prices to labor tax rates is poorly understood. Nonetheless, such elasticities play an important

[38] See, for example, Feldstein (1995), Saez (2010), and, for a nice review, Saez, Slemrod and Giertz (2012)

role in the formulas developed in Sections 1.4.1 and 1.4.2, suggesting further empirical investigations would be worthwhile. This could be dovetailed with a sufficient-statistic, rather than structural, approach to this sort of policy problem.

Finally, the present model assumes full commitment of the social planner. The various political economy constraints a government may face—especially in the context of an OLG economy—could lead to substantial changes to optimal policy and would be worth investigating seriously.

Chapter 2

Race and Home Price Appreciation in the United States: 1992–2012

2.1 Introduction

Different races lead vastly different economic lives in the United States. Semega, Fontenot and Kollar (2017) show a greater than $25,000 gap in median household income between white and black households in 2016, and this gap has been persistent since the passage of the Civil Rights Act.[1] The fact that such a gap persists despite legal efforts to close it has prompted much research into its causes.[2] Even more recently, as part of a broader trend toward focus on inequality, an effort has begun to document in much greater detail exactly the difficulties faced by minority groups. Chetty et al.

[1]For further documentation of the gap, see Harrison and Bennett (1995), Bound and Freeman (1992), Cain (1986), Kilbourne, England and Beron (1994), and many others.

[2]Reviews of this literature can be found in Chetty et al. (2018), Roland G. (2011), and most thoroughly Grodsky and Pager (2001). It is extensive and includes contributions on residential segregation (Wilson, 2012; Massey and Denton, 1993), classic prejudice on the basis of names (Bertrand and Mullainathan, 2004), poor achievement in school due to lack of same-race role models (Dee, 2005), and many others.

(2018) carefully explore the intergenerational transmission of wealth in different racial groups and find that, even conditional on parental income, black male descendants earn substantially less than their white male counterparts. They detail other results as well, such as astonishingly large differences in incarceration rates.

Related to this "income gap" is a "wealth gap." Various press articles (Jan, 2017; Jones, 2017; Shin, 2015) have commented on the gap's existence, and policy institutions have measured it to be around 5 times as large as the income gap (McKernan et al., 2013). The wealth gap has received substantial scholarly documentation as well (Smith, 1995; Oliver and Shapiro, 2006; Wolff, 1998). One of many possible reasons for such a gap is earning different returns on assets. Choudhury (2002) studies Health and Retirement Study data and finds substantial differences in investment choices across races. I focus on a different concern: Do blacks and whites earn different returns on the same asset class, and the largest part of most Americans' portfolios—their homes?

It is long documented that black neighborhoods are underpriced relative to white neighborhoods with the same characterisitics (Bailey, 1966; Harris, 1999; Cervero and Duncan, 2004), but that is a statement about the *level* of prices, not their *appreciation*. The popular press has posited that black households earn lower housing returns (Badger, 2016), but is this borne out in the data? And to what extent is any such gap related to differences in incomes or characteristics of the neighborhoods that different races select?

This paper sheds light on these and related questions by analyzing administrative data on over thirty million housing transactions, tied to race and income through a fuzzy link with administrative data on home mortgage applications, yielding over three million complete ownership spells with associated demographic data. It reveals a large and persistent gap of around 2.5 percentage points per year between the housing appreciations enjoyed by black and white households of the same income level. Accounting for differences in neighborhood racial makeup or changes thereto, home purchase price, and neighborhood income level reduces this effect to around one percentage point, but still leaves a statistically and economically significant gap. These magnitudes are quite large: they can lead to tens of thousands of dollars of lost wealth over a typical homeownership spell, and represent a substantial form of interracial economic inequality. These effects are qualitatively similar, though slightly quantitatively diminished, for Hispanics.

Along the way, I document large interracial differences in the attributes of the neighborhoods home buyers select. Black families, relative to white families of the same income, are more likely to live in black neighborhoods, neighborhoods that become even more black over time, poorer neighborhoods, and homes with a lower purchase price, most of which are in turn associated with lower home price appreciation. To give a sense of the magnitude of these differences, a typical black family in the *100th* percentile of a city's income distribution lives in a neighborhood that is more black than a white family in the *1st* income percentile. Similarly, a black family in the 75th percentile of income is likely to live in a neighborhood of similar income to that chosen by a typical white family in the 30th percentile of income.

Separating the data by commuting zone, I find that the extent of these interracial appreciation gaps has large geographical variation, with blacks faring much better on the coasts than in the middle of the country, and Hispanics faring better in a collection of cities that appear to have no common thread. Interestingly, these commuting zone level gaps are mostly unrelated to salient economic attributes of those cities, including the interracial mobility gaps documented by Chetty et al. (2018). This suggests that the appreciation gap is a phenomenon altogether different from other racial disparities documented in the literature.

Related Literature. This paper is not the first to investigate possible interracial disparities in home prices or their appreciation.[3] The former issue—differences in the *level* of home prices—based on neighborhood racial composition has received substantial attention as noted above, but is quite different from and more theoretically explicable than the latter, which focuses on *changes* in those home prices. The assertion that neighborhoods with a high percentage of minority residents have slow home price appreciation can be dated to at least the late 1970s, when Lake (1979) showed that suburban homes occupied by blacks were unlikely to be sold to whites, which could lead to priced disparities. Since then, a litany of researchers have studied various neighborhood-level aggregated (Kim, 2000; Anacker, 2010; Coate and Schwester, 2011) and survey (Long and Caudill, 1992; Coate and Vanderhoff, 1993; Flippen, 2004) data to test this hypothesis, usually finding that home price appreciation is in fact correlated with either the level or change in racial composition of a neighborhood in the case of neighborhood data, or the race of the homeowner in survey data. Other researchers have used a case study approach, focusing on particular neighborhoods and, often, superior, transaction-level data (Quercia et al., 2000; Macpherson and Sirmans, 2001; Coate and Schwester, 2011) to demonstrate different levels of home price appreciation for different races of homeowners. Bayer et al. (2013) analyze transaction-level data in four metropolitan areas and show that Hispanics overpay for homes, controlling for home and neighborhood characteristics.

To my knowledge, however, no paper has undertaken a recent, thorough, nationwide analysis of repeat sale home transactions, focusing on the demographics not only of the neighborhood but *of the home buyer himself*. As we will see, focusing on neighborhood attributes misses important parts of the story, and so the linkage between repeat sale transaction data and individual-level demographic data is an important contribution.

[3]See Flippen (2004) and Anacker (2010) for detailed reviews of the literature.

The rest of the paper is organized as follows. In Section 2.2, I discuss the many data sources compiled for this project and present summary statistics and maps of the availability of quality data. Section 2.3 documents the primary finding of the paper—large differences in home price appreciation experienced by homeowners of different races which cannot be explained by (even quite flexibly) controlling for income. Section 2.4 asks whether these gaps can be explained by differences in neighborhood characteristics selected by homeowners of different races but the same income, and shows that, while households of different races do indeed choose systematically different sorts of neighborhoods and homes, such differences cannot fully account for the different appreciation rates they experience. Finally, Section 2.5 discusses regional heterogeneity in the strength of the appreciation gap and shows that it is almost completely uncorrelated with salient economic attributes of the region, and Section 2.6 concludes.

2.2 Data

Two major microdata sources were used in this project: the DataQuick (DQ) database of home transaction and tax assessor records; and the Home Mortgage Disclosure Act (HMDA) Loan Application Register (LAR), which documents applicant demographics for most home mortgage applications made in the United States. Several smaller datasets were used in addition. In this section, I detail the relevant attributes of each dataset, relegating the most technical of details to Appendix B.1. I then discuss how the datasets were merged, and present availability and summary statistics for the merged data.

2.2.1 Data Sources

DataQuick. Most local governments (municipalities and counties) maintain records of properly reported real estate transactions and tax assessments of real estate performed within their jurisdictions. The DQ database of home transaction and tax assessor records is a proprietary database that aggregates this publicly available data in a consistent format (Ramirez et al., 2014). The tax assessment ("assessor") dataset documents physical attributes of properties within a jurisdiction, including exact physical location, mailing address, number of bedrooms and bathrooms, the presence of central air conditioning, year built, date of last permitted construction, and so on. The transaction history ("history") dataset details each transaction on these properties, including the transaction price, the names of the seller and buyer, what sort of transaction was performed (refinance, construction, resale, or subdivision), and any associated mortgage information (names of lenders and loan amounts). The history and assessor datasets are linked through a unique property identifier.

If a property i undergoes transactions, indexed by $t \in \{0, 1, 2, ...\}$, at dates $\{d_{i0}, d_{i1}, d_{i2}, ...\}$ (measured in days since a particular point in time) and for transaction price $\{p_{i0}, p_{i1}, p_{i2}, ...\}$, then I define annualized appreciation—the main outcome variable in this paper, as

$$A_{it} = \frac{\log p_{i,t+1} - \log p_{it}}{(d_{i,t+1} - d_{it})/365}.$$

.

While DQ standardizes the formats of each of these variables across a large variety of recorders' offices, it attempts very little data cleaning such as identification of duplicate transaction and flagging of implausible values. I mostly follow Guren (2018)[4] for data cleaning, but have made one major change that I discuss in Appendix B.1.

[4] I am grateful that Adam Guren left the cleaned data for future Harvard researchers.

In addition to issues with imperfect coverage, this dataset of course covers only homeowners, not renters.

Home Mortgage Disclosure Act Loan Application Register.

> Since 1976, most banks and other depository institutions...have been required under the HMDA to disclose to the public information about the geographic distribution of their loans for home purchase and home improvement...beginning in 1991, lenders for the first time have reported on all home loan applications they received and their disposition, plus the race or national origin, gender, and annual income of the applicants. (Canner et al., 1991)

This data is anonymized for public dissemination but importantly contains several variables that can be used to imperfectly match the data to the DQ data. The geographic location of the loan, up to the Census tract level, as well as the name of the lending institution and the loan amount, are disclosed, each of which has a corresponding variable in the DQ data.[5]

Smaller Datasets. Demographic data by block group was taken from the 1990, 2000, and 2010 Decennial Censuses, and covers the racial composition and income characteristics of Census block groups. Changes in variables are computed on a standardized set of geographies, so that they do not reflect composition bias. The geographical boundaries of the Census block groups are from Census TIGER/Line® Shapefiles.

Commuting zone attributes are drawn from the Equality of Opportunity Project (Chetty et al., 2014, 2018). The commuting zone delineations themselves are, in turn, from Autor and Dorn (2013).

[5]Government documents from the FFIEC website show that 1990 Census tracts were used from 1992 through 2002; 2000 Census tracts were used from 2003 through 2011, and 2010 Census tracts were used in 2012.

2.2.2 Linking and Cleaning

The most important link to be performed was between the DQ and HMDA data. The general idea is similar to Diamond and McQuade (2018). The DQ data has geographical, price, and loan information, but no demographic data. Meanwhile, the HMDA data has geographical, loan, and demographic information, but no price data. Data were linked by perfectly matching the Census tract,[6] loan amount in thousands, and first letter of the lending institution, and "fuzzily" matching the entire name of the lending institution using an expectation maximization algorithm (Sariyar and Borg, 2010) with a conservative cutoff.

The DQ data quality is quite poor, with many homes trading for implausibly low or high values. To combat this, as well as more standard concerns regarding results driven by outliers, I remove many observations. In particular, homes that have ever undergone construction (since the earliest observation) other than their initial building are dropped, along with any ownership spells lasting less than two years. Additionally, homes that ever were priced in the top or bottom 1% of their commuting zone-year are dropped, along with homes that ever had a reported price below \$10,000 or above \$20,000,000. Finally, I dropped ownership spells in which the home underwent annualized appreciation in the top or bottom 1% of spells in the commuting zone. Clerical review suggests that most of these spells were incorrectly recorded, and otherwise likely involved a highly unusual event.

2.2.3 Summary Statistics

The net result of all merging and cleaning is a dataset representing over thirty million transactions. Around 36% of those were successfully matched to demographic data from

[6] after correcting for changing Census boundaries—see Appendix B.1

the HMDA LAR. The low match rate is not particularly concerning; my data is quite large, as summary stats below will show, and so aggressive discarding of questionable data leads to very small loss of power and potentially substantial improvements in bias. Close to 30% of the transactions were followed by a subsequent (undropped) transaction on the same property, leaving approximately ten million complete homeownership spells.

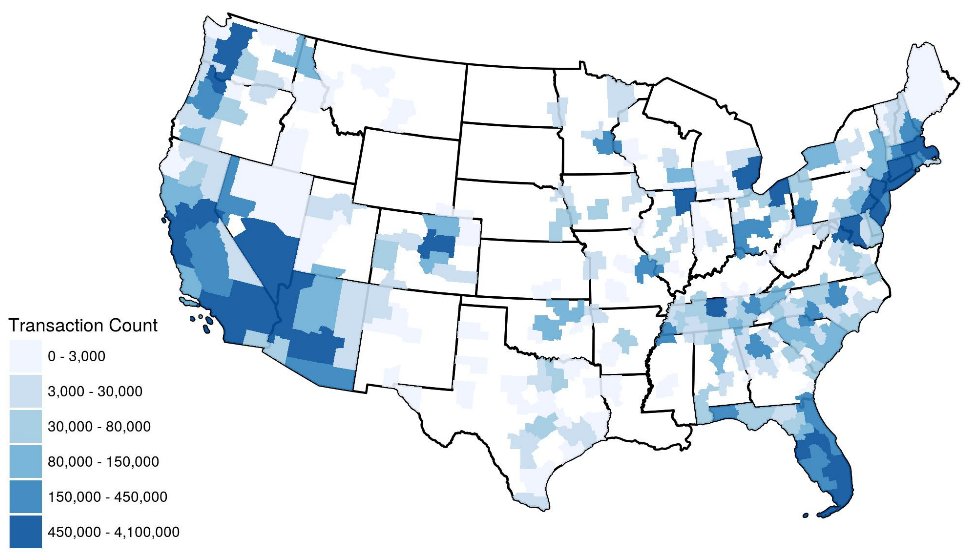


Figure 2.1: The count of transactions from the cleaned data by commuting zone. Outliers were dropped as in the text, but no requirements on matches to HMDA LAR data or Census data were enforced for this map.

Table 2.1 summarizes the full sample of transactions, while Figures 2.1 and 2.2 show the count of transactions by commuting zone, and the HMDA LAR match rate by commuting zone, respectively. Statistics are roughly in line with national averages, except that homes are substantially more expensive and households are substantially wealthier. This, in turn, is due to the overwhelming concentration of the non-dropped observations in the DQ data in wealthy coastal areas. Texas, in particular, lacks a

Table 2.1: Summary statistics for the full, cleaned sample of transactions

Variable	N	Mean	SD	Min	P25	Median	P75	Max
Price ($1000s)	33,364,910	218.343	190.731	10.001	100	165	275	15,850
Ann. Appreciation	9,843,953	0.037	0.119	-0.758	-0.009	0.042	0.101	0.935
Spell Length (Years)	9,843,953	6.124	3.612	2.003	3.425	5.121	7.767	24.474
Income ($1000s)	12,153,126	92.953	116.386	0	46	70	108	9,999
White	12,526,136	0.638	0.481	0	0	1	1	1
Asian	12,526,136	0.064	0.245	0	0	0	0	1
Black	12,526,136	0.063	0.244	0	0	0	0	1
Hispanic	12,526,136	0.125	0.331	0	0	0	0	1
Opp. Sex Couple	12,526,132	0.441	0.497	0	0	0	1	1
Single Male	12,526,132	0.179	0.383	0	0	0	0	1
Single Female	12,526,132	0.129	0.335	0	0	0	0	1
Male Couple	12,526,132	0.017	0.129	0	0	0	0	1
Female Couple	12,526,132	0.012	0.109	0	0	0	0	1
BG % Black (2000)	33,332,891	8.737	17.033	0	0.827	2.457	7.577	100
BG % Hisp (2000)	33,332,891	12.845	17.939	0	1.937	5.466	15.095	100
BG Med. HH Income ($1000s)	33,340,122	54.567	23.427	0	38.056	50.625	66.250	200.001
Change BG % Black (1990–2010)	33,035,778	2.985	10.361	-100	0	0.966	3.859	100
Change BG % Hisp (1990–2010)	33,035,778	8.520	11.736	-100	1.609	4.454	12.104	100

These statistics summarize the full sample of transactions after basic cleaning delineated in the text. Block group attributes in levels are measured within 2000 Decennial Census geographies, while block group attributes in changes are measured within 2010 Decennial Census geographies. "Hispanic" represents the "Hispanic" racial designation prior to 2004, and "Hispanic" ethnic designation combined with "white" racial designation from 2004 onwards. Other races represent the race in question prior to 2004, and represent the race in question, combined with "not Hispanic" ethnic designation from 2004 onwards.

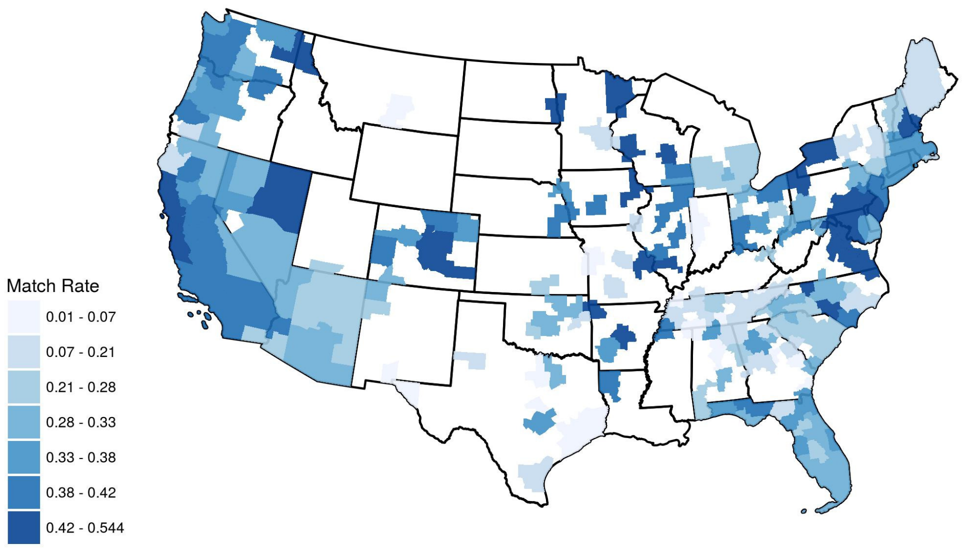


Figure 2.2: The match rate to HMDA LAR data by commuting zone. After outliers were dropped as in the text, the fraction of observations with a non-missing income (a variable from the HMDA LAR data that is never missing within that dataset) is recorded.

substantial number of observations, since most transactions in that state are mistakenly attributed a price of zero. Match rates are fairly uniform across the country, with somewhat better matching in most densely populated regions, and particularly poor matching in Texas, rural Tennessee, and Kansas City.

For the rest of the paper, analysis will be confined to the set of complete cases from this full sample. Complete cases are defined as those observations which have no missing variables. This will occur when the entire ownership spell, including ending transaction, is observed and matched to both HMDA LAR data and Census data. Additionally, commuting zones with less than 1000 such complete cases are dropped altogether, as conclusions should probably not be drawn from such a small collection of data.

Table 2.2 summarizes the sample of complete cases. Even this aggressive pruning of the dataset yields over three million cases with no missing variables—a very large sample for conducting the sorts of analyses to follow. The restriction to complete cases has changed the summary statistics very little, though spell lengths and annualized appreciations are reduced due to greater weighting toward the later part of the sample period, when data is more complete. Figure 2.3 shows the geographic distribution of these complete cases; it is apparent that they follow roughly the same distribution as the original counts.

2.3 The Appreciation Gap

In this section I document the main effect of the paper—a large, consistent, and nationwide gap between annualized home value appreciations earned by homeowners of different races, even after accounting for income differences. It is worth keeping in mind throughout that this effect is much more surprising than a result merely about home values; it loosely says something about the divergence of those values over time, even as

Table 2.2: Summary statistics for the sample of complete cases.

Variable	N	Mean	SD	Min	P25	Median	P75	Max
Price ($1000s)	3,182,339	228.935	173.839	10.200	116.500	175	285	3,300
Ann. Appreciation	3,182,339	0.026	0.124	-0.757	-0.027	0.042	0.103	0.830
Spell Length (Years)	3,182,339	5.413	2.898	2.003	3.230	4.679	6.775	21.405
Income ($1000s)	3,182,339	86.772	114.240	0	44	66	100	9,999
White	3,182,339	0.642	0.479	0	0	1	1	1
Asian	3,182,339	0.056	0.231	0	0	0	0	1
Black	3,182,339	0.054	0.225	0	0	0	0	1
Hispanic	3,182,339	0.143	0.350	0	0	0	0	1
Opp. Sex Couple	3,182,339	0.449	0.497	0	0	0	1	1
Single Male	3,182,339	0.109	0.311	0	0	0	0	1
Single Female	3,182,339	0.072	0.259	0	0	0	0	1
Male Couple	3,182,339	0.019	0.137	0	0	0	0	1
Female Couple	3,182,339	0.012	0.110	0	0	0	0	1
BG % Black (2000)	3,182,339	7.605	14.405	0	0.945	2.564	7.110	100
BG % Hisp (2000)	3,182,339	15.141	18.491	0	2.888	7.893	19.357	100
BG Med. HH Income ($1000s)	3,182,339	56.315	23.180	0	40.104	52.165	68.028	200.001
Change BG % Black (1990–2010)	3,182,339	2.617	9.398	-98.948	-0.020	0.966	3.585	100
Change BG % Hisp (1990–2010)	3,182,339	10.036	12.461	-99.103	2.016	5.860	14.854	81.895

These statistics summarize the sample of complete cases—cases that have no missing variables, and which remain after commuting zones with less than 1000 such cases were dropped altogether. Block group attributes in levels are measured within 2000 Decennial Census geographies, while block group attributes in changes are measured within 2010 Decennial Census geographies. "Hispanic" represents the "Hispanic" racial designation prior to 2004, and "Hispanic" ethnic designation combined with "white" racial designation from 2004 onwards. Other races represent the race in question prior to 2004, and represent the race in question, combined with "not Hispanic" ethnic designation from 2004 onwards.

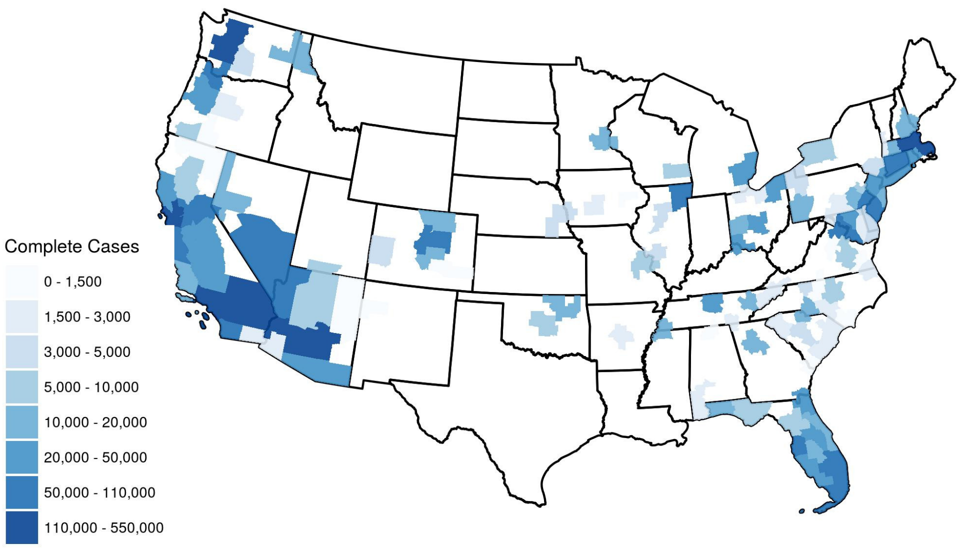


Figure 2.3: The number of complete cases by commuting zone. After outliers were dropped as in the text, only observations with non-missing annualized appreciation (i.e., an ending transaction date and price) and matches to the HMDA LAR data and Census demographic data were kept.

Chetty et al. (2018) documents the fact that racial income gaps have mostly stabilized. In the following sections, I will consider possible explanations for this phenomenon, but at present, I merely document it.

Though I have already mentioned the importance of accounting for income differences across races, there are other factors that must be controlled for when doing a proper analysis. First, one must account for the commuting zone of the home. This is important since, over the sample period, different regions of the U.S. experienced extremely different home price appreciations, and this is likely correlated with, but not caused by, the racial composition of the region. For example, California and the Eastern Seaboard have done very well over the last several decades, and these are also regions with fewer black residents.

Second, one must account for when the home was bought and sold. The sample period involves three periods of very different appreciations: pre-1996, 1996–2007 (the "boom"), and 2007–2012 (the "crash"). If black homeowners, for example, were much more likely to have purchased their homes in 2007 than 2000, for example, they will seem to have endured much worse home price appreciation, but that would be a spurious effect.

To account for these issues, for each of the subsequent graphs, I place all observations into cells according to the commuting zone, the year of purchase, and the year of sale (henceforth simply "cells"). Within each cell, I then demean annualized appreciation (though I add back the graph's overall mean to achieve the correct scale), and assign each observation an income percentile. Importantly, this is done *before* splitting by race. Thus, a 50th percentile black homeowner earns approximately the same income as a 50th percentile white homeowner. This allows for the comparison of incomes across commuting zones and times, even if those income distributions differ in moments other than the first.[7]

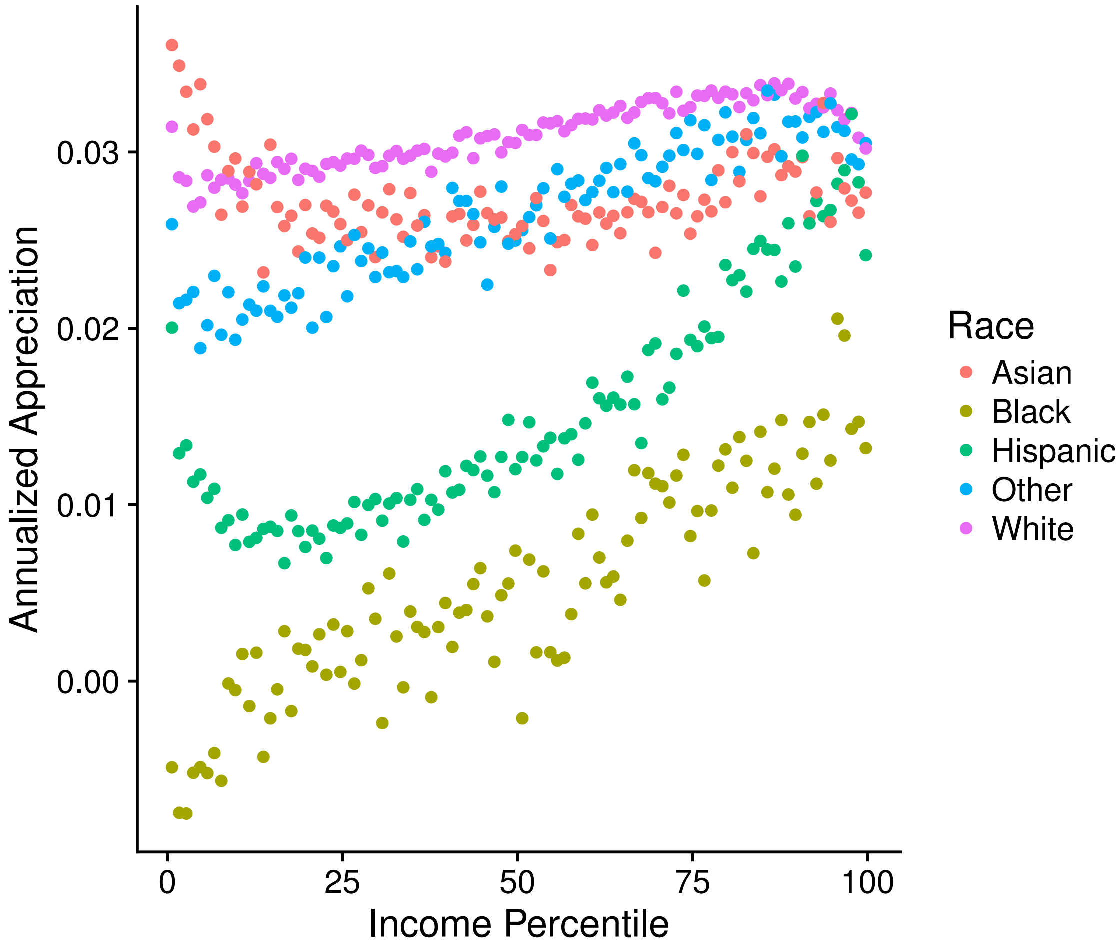


Figure 2.4: Binscatter of homeowner's income percentile versus annualized home price appreciation; points are colored according to the homeowner's race. Income percentiles are assigned within a commuting zone × purchase year × selling year cell, and appreciation values are demeaned within each cell; the overall mean is then added back to ensure the correct scale. Importantly, percentiles are assigned, and observations are binned along the horizontal axis, prior to splitting by race.

Figure 2.4 shows a binscatter[8] of these income percentiles versus these demeaned

[7] I did also experiment with simply demeaning incomes, but these led to far noisier graphs, suggesting that other moments do vary importantly, and that percentile is a much better predictor.

[8] After assigning the percentiles as above, observations are binned into one hundred equal-sized bins along the horizontal axis *prior* to splitting by race. Then, within each bin-by-race combination, averages of all observations' x and y values are taken, and the point is plotted. This is the process

appreciations. Overall, there appears to be a positive relationship between income and appreciation. However, this effect is completely dominated by the gap between races. Whites typically earn the largest returns, followed closely by Asians, and then much more distantly by Hispanics and blacks. Most of the Hispanic-white gap is closed for high earners, but the same cannot be said for blacks. The magnitude of this gap is quite large. There is a roughly 2.5 percentage point gap between returns earned by blacks and whites of median income; at the mean home value of approximately \$230,000, this represents \$5,750 of lost earnings *per year, compounded*.

Figure 2.5 breaks out this graph into homeownership spells that occurred entirely within the real estate boom of 1996–2007 and those that occurred entirely during or after the ensuing crash. Such separation reveals heterogeneity in the effect across economic conditions which is masked by the pooled graph. In particular, the poorer households fared substantially better than wealthier households during the boom, but substantially worse during the crash; this is consistent with conventional accounts of the real estate bubble and its bursting.

More interestingly, the interracial gaps are much larger during the crash than the boom. Hispanics' returns were very similar to whites' returns during the boom—even substantially exceeding those returns for poorer households—but trailed whites' returns during the crash by around 4 percentage points. Poor blacks' returns continued to lag those of their white counterparts by around 2.5 percent points during the boom, though that gap closes substantially for higher income blacks; more importantly, the crash appears to have been particularly disastrous for blacks of all incomes, with a black-white return gap of around 7 percentage points throughout the income distribution.

---

followed for all future binscatters unless stated otherwise.

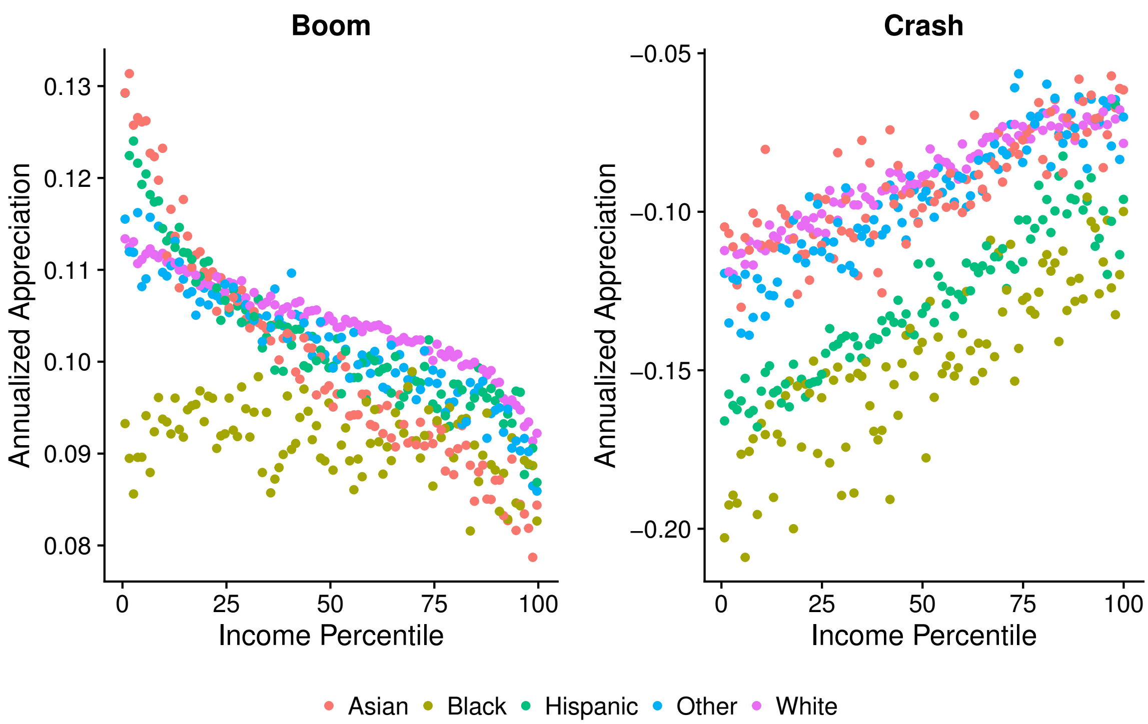


Figure 2.5: Binscatter of homeowner's income percentile versus annualized home price appreciation, broken out by period of homeownership; points are colored according to the homeowner's race. The left panel represents homeownership spells during the real estate boom of 1996–2007, while the right panel represents homeownership spells during or after the crash of 2007. Income percentiles are assigned within a commuting zone × purchase year × selling year cell, and appreciation values are demeaned within each cell; the overall mean is then added back to ensure the correct scale. Percentiles are assigned, and observations are binned along the horizontal axis, prior to splitting by race.

Table 2.3: Regressions using only race and income percentiles as explanatory variables

	Dependent variable:		
	Annualized Appreciation		
	Full Sample	Boom	Crash
	(1)	(2)	(3)
Asian	-0.538^{***}	-0.495^{***}	-0.754^{*}
	(0.122)	(0.137)	(0.421)
Black	-2.769^{***}	-1.342^{***}	-6.553^{***}
	(0.259)	(0.195)	(0.582)
Hispanic	-1.940^{***}	-0.267	-4.869^{***}
	(0.185)	(0.166)	(0.479)
Other	-0.540^{***}	-0.263^{***}	-0.888^{***}
	(0.059)	(0.051)	(0.118)
Income Percentile	0.005^{**}	-0.017^{***}	0.048^{***}
	(0.002)	(0.006)	(0.007)
Income Percentile × Asian	-0.006^{**}	-0.022^{***}	-0.0003
	(0.003)	(0.003)	(0.006)
Income Percentile × Black	0.014^{***}	0.014^{***}	0.028^{***}
	(0.004)	(0.005)	(0.009)
Income Percentile × Hispanic	0.009^{***}	-0.014^{***}	0.023^{***}
	(0.004)	(0.003)	(0.006)
Income Percentile × Other	0.008^{***}	-0.005^{**}	0.016^{***}
	(0.001)	(0.002)	(0.003)
Observations	3,182,339	1,474,849	135,053
R^2	0.716	0.413	0.494
Adjusted R^2	0.715	0.411	0.490
Note:			*p<0.1; **p<0.05; ***p<0.01

These regressions estimate (2.1) and summarize the relationship between race and home price appreciation, controlling for income percentile. Income percentiles are calculated within a purchase year-selling year-commuting zone cell to allow for comparison across these cells. The dependent variable is annualized home price appreciation for all models. All models have cell-fixed effects. The equation is estimated on three separate samples—the full sample, the "boom" sample, and the "crash" sample, respectively. Standard errors (in parentheses) are clustered at the commuting zone level and calculated via the block bootstrap. All coefficients should be interpreted as percentage points.

Table 2.3 reports the results of analogous regressions, which estimate

$$A_{ic} = \alpha_c + \beta_0 Y_{ic} + \sum_{r \neq \text{White}} \alpha_r R_i^r + \sum_{r \neq \text{White}} \beta_r R_i^r Y_{ic} + \epsilon_i, \tag{2.1}$$

where i indexes observations, c indexes cells, Y_{ic} is the income percentile of observation i within cell c[9], R^r are a set of race dummies, and A is annualized appreciation. Standard errors are clustered at the commuting zone level to account for cross-sectional correlation in error terms within a commuting zone, as well as autocorrelation of the error term. Due to extremely unequally sized clusters, they are calculated through the block bootstrap, adjusted according to Sherman and le Cessie (1997). Column 1 carries the estimates for the whole dataset, while columns 2 and 3 focus on the boom and crash periods, respectively.

Flexible Controls. Three problems exist with this manner of controlling for income, and they push in different directions. First, percentile may not be the most relevant statistic. It seems quite possible that both income percentile and absolute deviation from the mean log income may predict a household's fortunes in the housing market, possibly in a nonlinear way. More flexibly controlling for income may eliminate some of the racial gap observed in Figures 2.4 and 2.5. Second, the process that maps income into appreciation may differ substantially from commuting zone to commuting zone. The regressions and graphs above allow for this only through differing income distributions, by using percentiles. If, for example, blacks are disproportionately poor, and disproportionately live in areas in which poor residents fare badly in the housing market, then the above graphs and regressions may overstate the "true" racial gap. Finally, the outcome variable is demeaned within each cell, which could be a form of

[9]demeaned across the dataset as a whole to allow interpretation of fixed effects as gaps, relative to whites, at the median income level

"overfitting"; even if some cells' "true" population means are the same, this process will overfit to the data actually observed.

All three of these concerns can be simultaneously addressed using a flexible machine learning approach.[10] The goal is to predict annualized appreciation using the time of purchase and sale, commuting zone, and income, but *not* race. In particular, I apply the method of gradient boosted trees (Friedman, 2001) to variables that have not been normalized or standardized in any way. I use the "LightGBM" package's popular implementation. This method led to better fit than the linear model above, even though it implicitly had to predict the *level* of appreciation within each cell and account for differences in income distributions across commuting zones, since the outcome was not cell-demeaned and income was not converted to percentiles.

Figure 2.6 plots a binscatter of the results of this exercise. On the horizontal axis is the algorithm's predicted appreciation, and the vertical axis is the error in the prediction. We see a strong pattern: there continues to be an interracial gap in appreciation, with that gap particularly large at times, places, and income levels with poor overall appreciation. This is consistent the "boom" versus "bust" story from Figure 2.5.

After calculating these predictions, I estimate the following equation:

$$A_i = \alpha + \beta_0 \tilde{A}_i^{\text{GBT}} + \sum_{r \neq \text{White}} \alpha_r R_i^r + \sum_{r \neq \text{White}} \beta_r R_i^r \tilde{A}_i^{\text{GBT}} + \epsilon_i, \tag{2.2}$$

where $\tilde{A}_i^{\text{GBT}}$ represents the gradient boosted trees prediction of annualized appreciation for observation i. Table 2.4 contains the results, which are qualitatively similar to those using just income percentile. Overall, they results strengthen the evidence for the overall effect: a strong interracial gap between home appreciation, with that gap

[10] For more details on these methods, see Appendix B.2

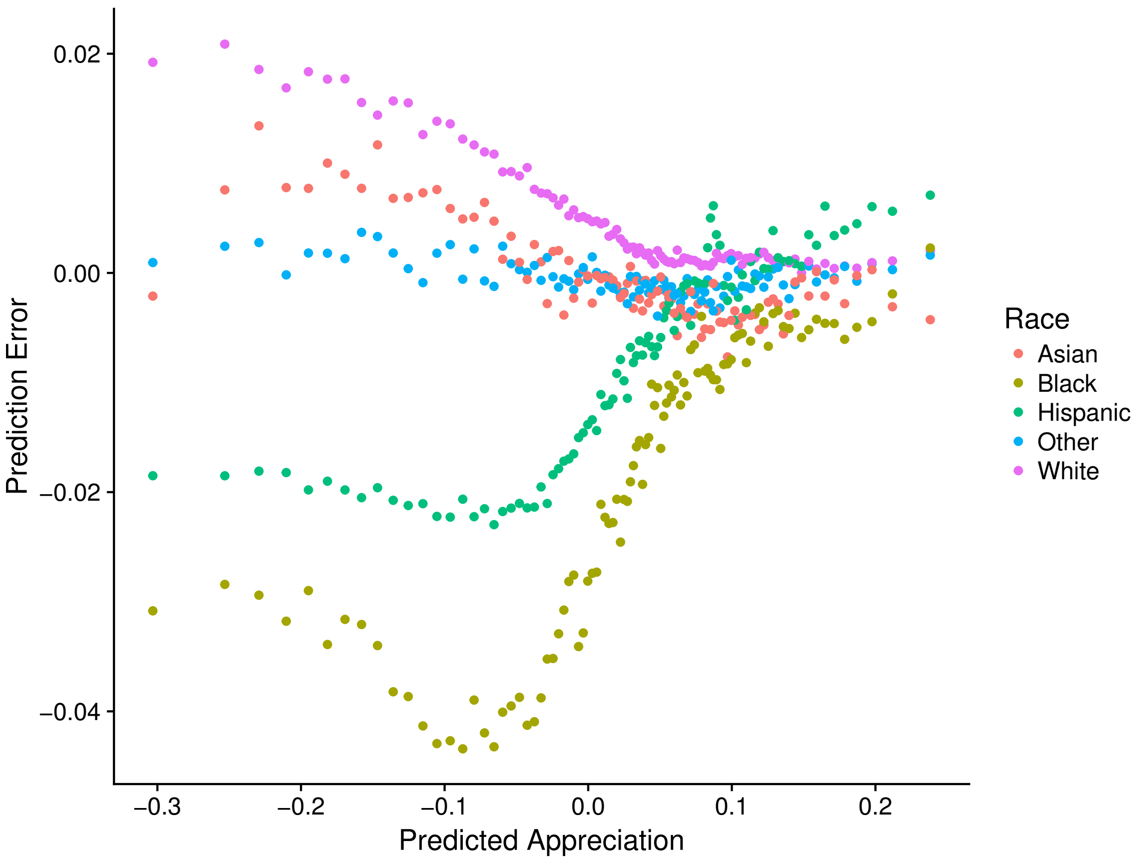


Figure 2.6: Binscatter of predicted home price appreciation versus the error in the prediction; points are colored by the homeowner's race. Predictions were performed using boosted trees with purchase date, sale date, commuting zone, and income as predictors of appreciation, with no normalization or standardization necessary. Observations are binned along the horizontal axis prior to splitting by race.

Table 2.4: Regression of realized home price appreciation on appreciation predicted by gradient boosted trees from income variables alone

	Dependent variable:
	Annualized Appreciation
Asian	-0.485^{***}
	(0.117)
Black	-2.389^{***}
	(0.183)
Hispanic	-1.252^{***}
	(0.149)
Other	-0.504^{***}
	(0.055)
Predicted Appreciation	0.953^{***}
	(0.080)
Predicted Appreciation × Asian	0.014
	(0.017)
Predicted Appreciation × Black	0.144^{***}
	(0.023)
Predicted Appreciation × Hispanic	0.110^{***}
	(0.018)
Predicted Appreciation × Other	0.040^{***}
	(0.006)
Observations	3,182,339
R^2	0.747
Adjusted R^2	0.747
Note:	*p<0.1; **p<0.05; ***p<0.01

This regression estimates (2.2) and summarizes the relationship between race and home price appreciation, controlling flexibly for any effects of income by predicting appreciation from income using gradient boosted trees. The dependent variable is annualized home price appreciation for all models. Predictions are performed using explanatory variables of purchase date, selling date, commuting zone, and income, with no standardization. Standard errors (in parentheses) are clustered at the commuting zone level and calculated via the block bootstrap; they account for uncertainty around the machine learning prediction. All coefficients should be interpreted as percentage points.

being felt especially strongly at times, places, and incomes that are associated with poor appreciation.

2.3.1 Risk

The most standard explanation for differing asset returns is different levels of risk. Consumption CAPMs identify risk as covariance with consumption, which in turn is often interpreted as covariance with the market return, or "beta." Applying a similar argument to the housing market is more theoretically problematic, as Sinai, Souleles and Slnai (2005) points out that home ownership can act as a hedge against future housing costs; in fact, the more the home's value covaries with the housing market, the more useful that hedge becomes, to a point. However, cell-demeaned appreciation is somewhat adjusted for covariance with the (commuting zone) housing market,[11] and so analyzing the variance of this demeaned quantity within the race-income bins above may give some insight into whether these return gaps can be explained by whites incurring greater risk than blacks or Hispanics.

Figure 2.7 shows a binscatter of the variance of annualized appreciation. This binscatter was created in precisely the same way as Figure 2.4, except that the standard deviation of appreciation within the observations represented by a point, rather than the mean, is plotted on the vertical axis. The result shows precisely the *opposite* of what might explain the appreciation gap; those portfolios of homes earning the *lowest*

[11] Any residual variance must either be correlated with the housing market, anticorrelated with the housing market, or orthogonal. If it is orthogonal, it represents standard risk as the housing cost hedging motive is absent. If it is correlated with the housing market, then the home is "too much of a hedge" against the market, leaving excessive risk that must be compensated with higher returns. If it is anticorrelated with the housing market, then any increase in this residual variance makes the home a "worse hedge", and therefore must be compensated with higher returns. This analysis does not consider every possible arrangement of the residual variance—it may have different properties in different parts of the distribution than others—but it suggests that the residual variance—the variance of these demeaned appreciations—should generally be compensated with higher returns.

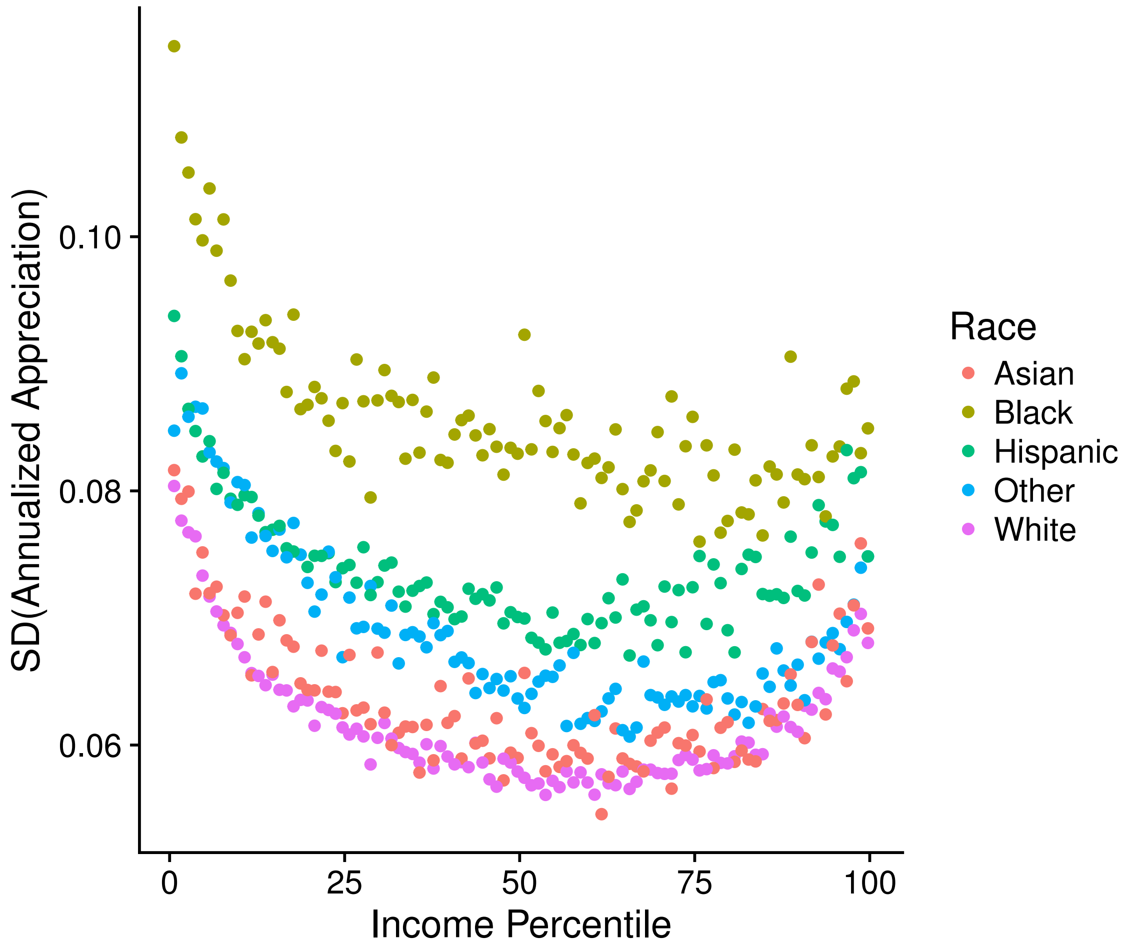


Figure 2.7: Binscatter of homeowner's income percentile versus standard deviation of annualized home price appreciation; points are colored by homeowner's race. Income percentiles are assigned within a commuting zone × purchase year × selling year cell, and appreciation values are demeaned within each cell. Percentiles are assigned, and observations are binned along the horizontal axis, prior to splitting by race. Finally, the standard deviation of appreciation is calculated within each bin × race, rather than the usual binscatter statistic–the mean.

returns—homes owned by blacks and Hispanics, and homes owned by low income households—are also those with the *highest* variance in annualized appreciation. This is consistent with Figure 2.5's breakdown of the racial return gap by period; blacks and Hispanics did much better during the boom than during the crash, which suggests that they are in fact incurring more (downside) risk than whites.[12] Figure 2.8 breaks

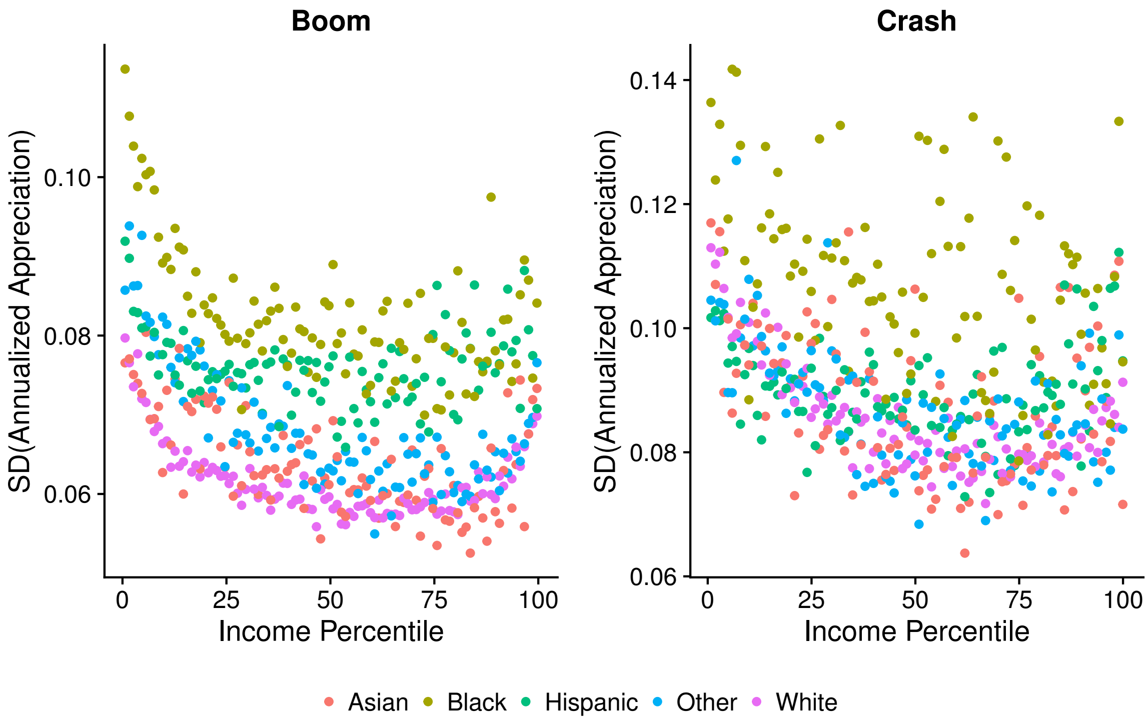


Figure 2.8: Binscatter of homeowner's income percentile versus standard deviation of annualized home price appreciation, broken out by period of homeownership; points are colored by homeowner's race. Income percentiles are assigned within a commuting zone × purchase year × selling year cell, and appreciation values are demeaned within each cell. Percentiles are assigned, and observations are binned along the horizontal axis, prior to splitting by race. Finally, the standard deviation of appreciation is calculated within each bin × race, rather than the usual binscatter statistic–the mean.

out this variance analysis by period. Here we see that the risk distribution across races and incomes is similar during the boom and crash. Since the interracial gap in housing returns is substantially different between the two periods, but the risk gap is

[12]though this sort of market-wide risk is eliminated for Figure 2.7 through the cell-demeaning process

not, this provides further evidence that a risk (or hedging) effect cannot account for the interracial gap in returns.

2.4 Possible Explanations

So far, I have only documented a large gap between home appreciations earned by whites and other races after flexibly accounting for differences in income distributions between the races. One possible explanation for such a gap would be differences in the sorts of houses purchased by households of different races, but the same incomes. In this section, I demonstrate precisely such differences, both in the price of the home, and the neighborhood characteristics; the presence of such difference are unlikely to surprise the reader, but the magnitudes, even after conditioning on income, are quite large. Despite such large differences, however, I will argue that they can explain a very small portion of the interracial appreciation gap discussed in the previous section.

2.4.1 Interracial Differences in Home Choices

Racial segregation seems to be a permanent fixture of life in the United States (Massey and Denton, 1993). It persists even among families at the top of the income distribution, and leads black households to live in poorer neighborhoods as well (Eligon and Gebeloff, 2016). In the ensuing graphs, I demonstrate the extent of this phenomenon along by plotting several neighborhood characeristics (racial makeup in 2000—roughly the midpoint of my sample period, change in the racial makeup from 1990 to 2010, median income in 2000) and the purchase price itself against income percentile, separated by race. As before, these binscatters involve income percentiles that are assigned within purchase year-selling year-commuting zone cells, and all of the homes' characteristics are demeaned within that cell as well.

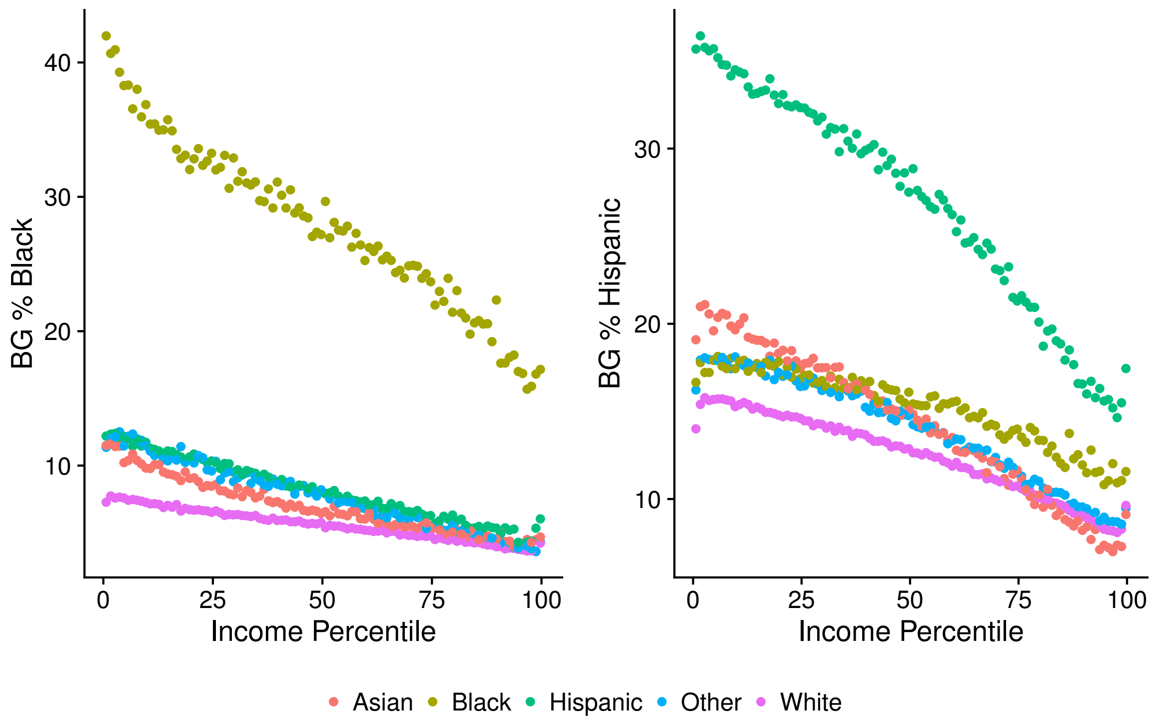


Figure 2.9: Binscatter of homeowner's income percentile versus racial makeup of the Census block group; points are colored by homeowner's race. The left panel uses the black percentage, while the right uses the Hispanic percentage, as the outcome variable. Income percentiles are assigned within a commuting zone × purchase year × selling year cell, and the variable on the vertical axis is demeaned within each cell; the overall mean is then added back to ensure the correct scale. Percentiles are assigned, and observations are binned along the horizontal axis, prior to splitting by race.

The first pair of such graphs, found in Figure 2.9 starkly demonstrate the extent of racial segregation across the county. Higher income households of all racial groups are less likely, relatively to their poorer counterparts, to choose neighborhoods with large populations of underrepresented minorities. However, the difference across races is striking. Black families in the 100th percentile are likely to choose a neighborhood that is substantially more black than that chosen by even the poorest percentile of white families. Similarly, Hispanic families in the 100th percentile of income are likely to choose a neighborhood that is roughly as Hispanic as that chosen by the poorest percentile of white families. Interestingly, there appears to be very little of a cross-effect; black families live in neighborhoods that are only slightly more Hispanic than white families of similar incomes, and Hispanic families live in neighborhoods that are only slightly more black than white families of similar incomes.

Figure 2.10 shows that the segregation effect exists not just in levels, but also in changes. For instance, black families live in neighborhoods that are unusually black, but also become even more black over time.[13] Furthermore, the same relationship with income exists: richer families choose neighborhoods that become more white over time, while poorer families choose neighborhoods that become more populated by underrepresented minorities. Again, the magnitudes of these interracial gaps are quite large.

Figure 2.11 considers the wealth of the neighborhood, as manifest in the neighborhood's median household income and the purchase price of the home itself. It is unsurprising that wealthier households choose wealthier neighborhoods and more expensive homes, but once again, the interracial gaps are quite large. For instance, a black family at the 75th percentile of the income distribution chooses a neighborhood that is similarly wealthy as that chosen by a white family at around the 30th percentile

[13]Some of that effect is of course mechanical.

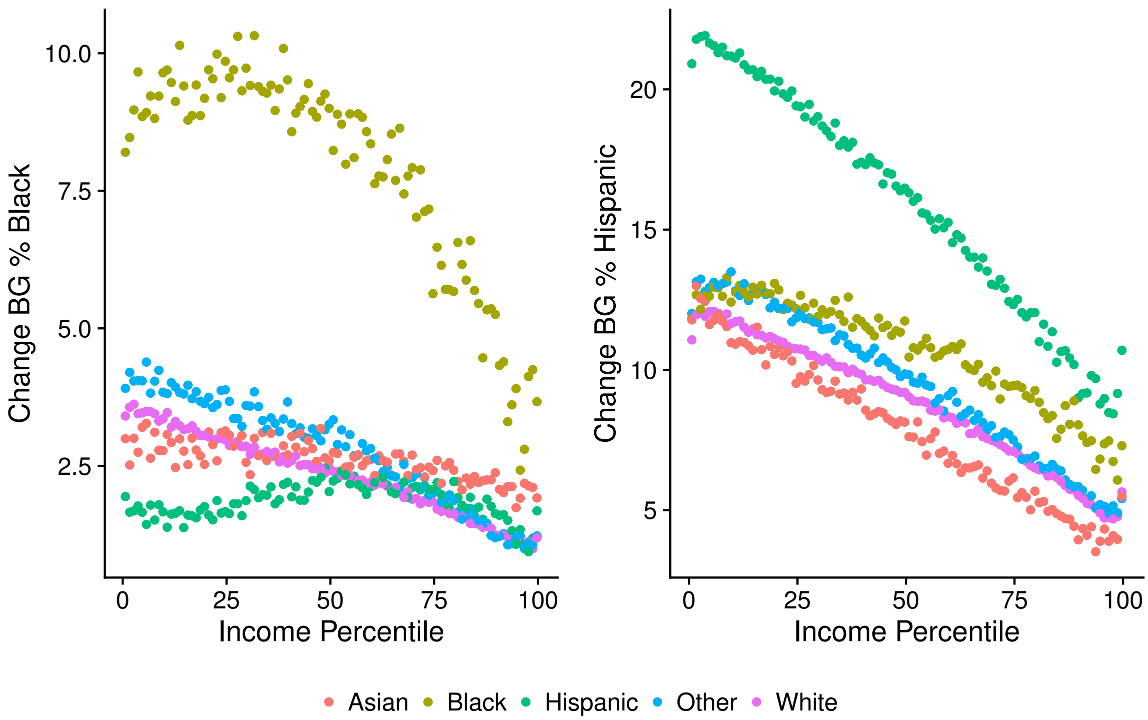


Figure 2.10: Binscatter of homeowner's income percentile versus change in the racial makeup of the Census block group; points are colored by homeowner's race. The left panel uses the change in the black percentage, while the right uses the change in the Hispanic percentage, as the outcome variable. Income percentiles are assigned within a commuting zone × purchase year × selling year cell, and the variable on the vertical axis is demeaned within each cell; the overall mean is then added back to ensure the correct scale. Percentiles are assigned, and observations are binned along the horizontal axis, prior to splitting by race.

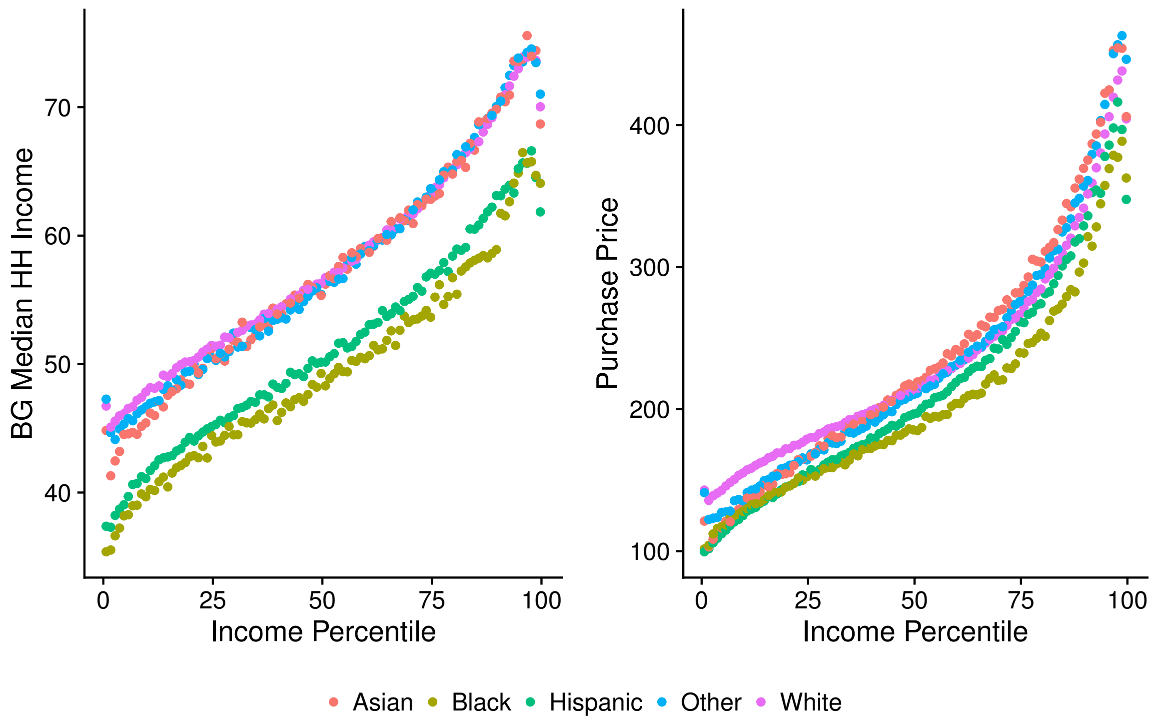


Figure 2.11: Binscatter of homeowner's income percentile versus change in the racial makeup of the Census block group; points are colored by homeowner's race. The left panel uses the change in the black percentage, while the right uses the change in the Hispanic percentage, as the outcome variable. Income percentiles are assigned within a commuting zone $\times$ purchase year $\times$ selling year cell, and the variable on the vertical axis is demeaned within each cell; the overall mean is then added back to ensure the correct scale. Percentiles are assigned, and observations are binned along the horizontal axis, prior to splitting by race.

of the distribution. Similarly, wealthy black families purchase homes for lower prices than their white counterparts, though that gap is somewhat smaller.

2.4.2 Effect of Neighborhood Attributes on Appreciation

Having seen that underrepresented minorities choose neighborhoods with substantially different characteristics than their white counterparts of similar income levels, it seems worthy of investigation whether this might explain the appreciation gap. In particular, these neighborhood attributes might well be correlated with home price appreciation, and one might ask whether underrepresented minorities—controlling for neighborhood attributes rather than income—still experience worse home price appreciation. We will see that they do.

A first attempt at answering this question would recreate Figure 2.4, except replacing income percentile on the horizontal axis with the percentile of the neighborhood attribute in question, as it eliminates covariance between these attributes and race of the home buyer. Figure 2.12 presents precisely such graphs. In most cases, it is clear that the neighborhood attribute is correlated with appreciation in a way that would explain part of the interracial appreciation gap. For example, more black neighborhoods are more likely to have lower appreciation, as are neighborhoods that become more black over time, and neighborhoods that are poorer. The predictive value is not particularly strong, but that is not surprising since, if these neighborhood attributes were strongly correlated with appreciation, then an arbitrage opportunity would exist.

Perhaps more noticeable is a persistent interracial gap across all six graphs. In all graphs, the typical gap between black and white appreciations for a given type of neighborhood is roughly the 2.5 percentage point gap that we saw in Section 2.3. This suggests that the differing neighborhood characteristics chosen by different races are not solely responsible for the gap observed in Section 2.3.

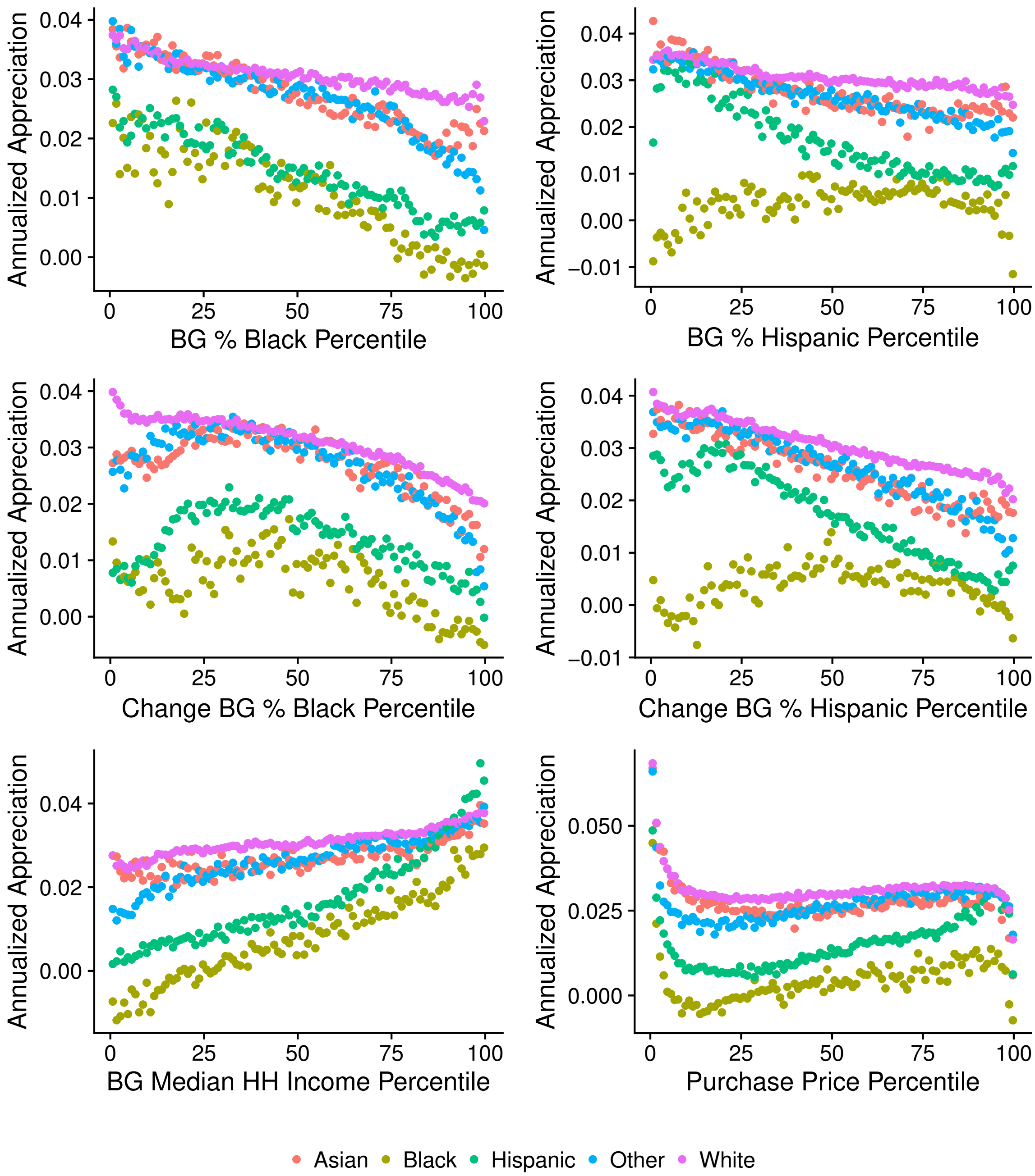


Figure 2.12: Binscatter of annualized appreciation versus various attributes of the neighborhood of the home (plus the home's purchase price); points are colored by homeowner's race. Each of these attributes are phrased in terms of percentiles within a commuting zone $\times$ purchase year $\times$ selling year cell, and annualized appreciation is demeaned within each cell; the overall mean is then added back to ensure the correct scale. Percentiles are assigned, and observations are binned along the horizontal axis, prior to splitting by race.

However, it is only suggestive evidence, as any one of these graphs will miss any contribution from the other five characteristics. For example, black families in very black neighborhoods may fare worse than white families in similarly black neighborhoods if they choose poorer black neighborhoods than the corresponding white families; the graph in the lower left hand corner suggests that richer neighborhoods fare better. More broadly, these graphs fail to pick up on any interaction between these characteristics and/or the buyer's income; it could be that flexibly controlling for *all* of these characteristics is sufficient to explain the interracial gap.

To test this, I return to the machine learning methods of the last section. Gradient boosted trees method will be applied to the all six of the characteristics in Figure 2.12, plus buyer income, purchase date, selling date, and commuting zone, with no normalization necessary.

The left panel Figure 2.13 plots binscatters of the prediction error versus predicted appreciation, broken out by the homeowner's race. We see that the gap remains, although it is clearly smaller than when controlling for only income. This suggests that these neighborhood factors, taken together, account for a large portion of the gap documented in Section 2.3. The first column of Table 2.5 estimates (2.2) using these new predictions. We see that the racial gaps remain both economically and statistically significant, although they have dropped sharply—for example, to around 0.8 percentage points for the black-white gap. This suggests that a large portion of the racial gaps are due to differential neighborhood selection, even conditional on income.

One concern with these predictions is that they may implicitly account for the race of the homeowner, even though they were not formed using that variable. The reason is that many of the predictor variables are strongly correlated with homeowner race, as seen in Figures 2.9 through 2.11. Even if, conditional on homeowner race, these demographics do not predict home price appreciation, a machine learning algorithm will

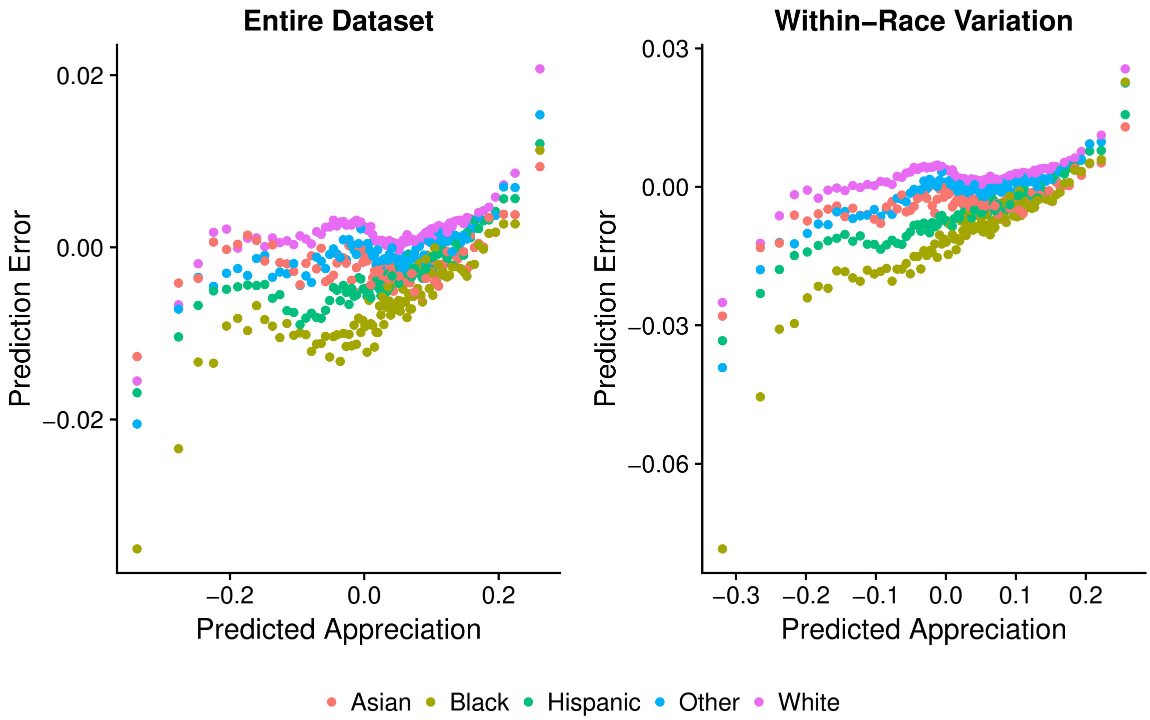


Figure 2.13: Binscatter of predicted home price appreciation versus the error in the prediction; points are colored by the homeowner's race. To make the predictions I use boosted trees with purchase date, sale date, commuting zone, income, and the 6 characteristics in Figure 2.12 as predictors of appreciation, with no normalization or standardization necessary. Observations are binned along the horizontal axis prior to splitting by race.

Table 2.5: Regressions of realized home price appreciation on appreciation predicted by gradient boosted trees from income and neighborhood characteristics.

	Entire Dataset	Within-Race Variation
	(1)	(2)
Asian	-0.351^{***}	-0.554^{***}
	(0.040)	(0.071)
Black	-0.800^{***}	-1.435^{***}
	(0.080)	(0.081)
Hispanic	-0.412^{***}	-0.811^{***}
	(0.053)	(0.101)
Other	-0.201^{***}	-0.300^{***}
	(0.021)	(0.032)
Predicted Appreciation	1.014^{***}	1.021^{***}
	(0.095)	(0.094)
Predicted Appreciation × Asian	-0.007	-0.013^{**}
	(0.006)	(0.006)
Predicted Appreciation × Black	0.029^{***}	0.018^{*}
	(0.007)	(0.009)
Predicted Appreciation × Hispanic	0.013^{***}	0.021^{***}
	(0.004)	(0.008)
Predicted Appreciation × Other	0.008^{***}	0.005
	(0.002)	(0.004)
Observations	3,182,339	3,182,339
R^2	0.835	0.821
Adjusted R^2	0.835	0.821
Note:		*p<0.1; **p<0.05; ***p<0.01

These regressions estimate (2.2) and summarize the relationship between race and home price appreciation, controlling flexibly for income plus all characteristics in Figure 2.12. The dependent variable is annualized home price appreciation for both models. Both columns use gradient boosted trees to predict appreciation from purchase date, selling date, commuting zone, and all characteristics from Figure 2.12. The first column applies the algorithm to the entire dataset; the latter trains the model in a manner that only-uses intra-racial variation (see text). Predictions are demeaned, such that race-fixed effects should be interpreted in appreciation gaps at the mean level of predicted appreciation. Standard errors (in parentheses) are clustered at the commuting zone level and calculated via the block bootstrap; they account for uncertainty around the machine learning prediction. All coefficients should be interpreted as percentage points.

use them to implicitly predict the homeowner's race, and therefore mistakenly attribute a "true" homeowner race effect to these demographic variables. While regularization[14] prevents overfitting, it does *not* protect against this sort of Bayesian updating.

This concern is not unique to the present case; it is similar to concerns about machine learning methods implicitly discriminating on the basis of race, even if they are not provided race as a predictor (Miller, 2015). Hardt, Price and Srebro (2016) attempt to derive methods for optimally revising a prediction to make sure that it satisfies various intuitive properties of fairness. However, their methods all include using the prohibited predictor (race) in the adjustment process, which for the present context seems inappropriate.

To combat this problem, I use a method that exploits only within-race variation in home appreciation to make predictions. In particular, I separately train the algorithm on separate subsets of the data—one for each race—and then use each such algorithm to make predictions for the entire dataset. For example, when I use the version trained only homeownership spells with white homeowners, I am implicitly predicting the appreciation of that homeownership spell, controlling for neighborhood characteristics and income, but *assuming the homeowner is white.* I then take the population-weighted average of these predictions to construct the final prediction.[15] This prevents any use of covariance between predictors—like neighborhood characteristics and income—and race in forming the predictions.

The results of this exercise can be found in the right panel of Figure 2.13 and the second column of Table 2.5. Qualitatively, they are similar to the results found in the left panel and the first column, respectively, but the interracial differences are

[14] See Appendix B.2 for more details.

[15] That is, if only two races, A and B are found in the data, each with an overall 50% share, I take the unweighted mean of the prediction formed by training the algorithm only on A observations, and the prediction formed by training the algorithm only on B observations.

much larger. This suggests that the previous conclusion that different neighborhood characteristics account for much of the interracial gap should be tempered; when using only within-race variation, it appears that blacks and Hispanics do fare substantially worse than their white counterparts, even when selecting neighborhoods with similar characteristics.

2.4.3 Other Possible Explanations

What could possibly account for a substantial interracial gap in appreciation, even conditional on neighborhood characteristics and income? At least four explanations are possible, but none of these can be tested in the present context.

First, the observed characteristics certainly do not cover all aspects of neighborhoods that affect its appreciation. School quality, crime rate, pollution level, proximity to amenities, and many other characteristics would be relevant. To the extent that these other factors are not correlated with the observed factors, but are correlated with the race of the homeowner, these may explain part of the appreciation gap.

Second, even if the observed characteristics capture most relevant information about a neighborhood, they are necessarily tabulated in discrete geographical areas. The lines between these areas may not always be drawn in a manner consistent with the local residents' idea of the neighborhood; one might need to "zoom out," "zoom in," or generally shift the boundaries to capture the "true" neighborhood characteristics.[16]

Third, minority home buyers may overpay for their homes, and/or minority home sellers may underprice their homes. In fact, Bayer et al. (2013) document precisely such a phenomenon for Hispanics in four cities, showing that they tend to overpay by around 3%. While this helps explain that gap, it can only fully account for the 0.8 percentage

[16]In future work, I will consider robustness to using tract characteristics, which is equivalent to "zooming out."

point per year gap found in the second column of Table 2.5 if homeownership spells tend to range between three or four years; otherwise, this three percent gap would be amortized over too long a period.

Finally, the data only imperfectly accounts for investment in the home during a homeownership spell. While I have dropped all homes that underwent permitted construction, this does not account for relatively minor projects that do not require permits, or more general upkeep. There is evidence that white homeowners plan more remodeling than other races (The Harvard Joint Center for Housing Studies, 2015), which could explain some of the gap.

2.5 Regional Differences

The foregoing sections have identified a large gap in home price appreciations experienced by homeowners of different races in the United States. In this section, I investigate whether there are regional differences in this effect, and whether any such differences are correlated with salient economic attributes.

To do so, I estimate 2.2 separately for each commuting zone, with predictions formed only based on income (analogous to Table 2.4) and with predictions formed based on many characteristics (analogous to Table 2.5; I consider estimates of the variety similar to column 2, using only intra-racial variation in characteristics).[17] I then tabulate the coefficients from that estimation; they represent, in a given commuting zone, the interracial gaps in home price appreciation, controlling for other factors.

Figure 2.14 shows the results of this exercise for the black-white appreciation gap, while Figure 2.15 does the same for the Hispanic-white gap. In each case, the top

[17] Predictions are *not* performed on a city-by-city basis, as this would lead to a high amount of overfitting.

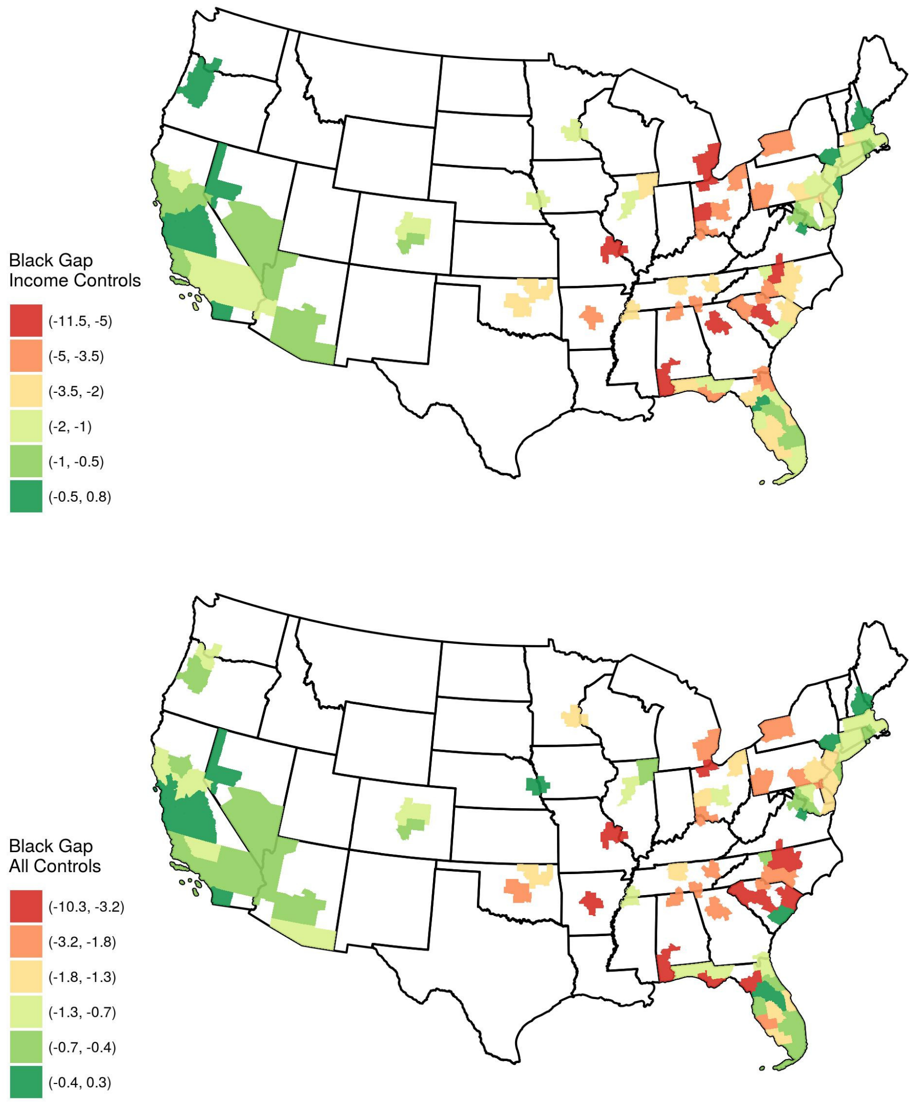


Figure 2.14: Map of the black-white appreciation gap by commuting zone. Areas in red have particularly problematic gaps, while areas in green have much less problematic—or even reversed—gaps. The top panel measures gaps relative to predictions based only on income and date, while the bottom panel measures gaps relative to predictions based on all variables discussed in Section 2.4.

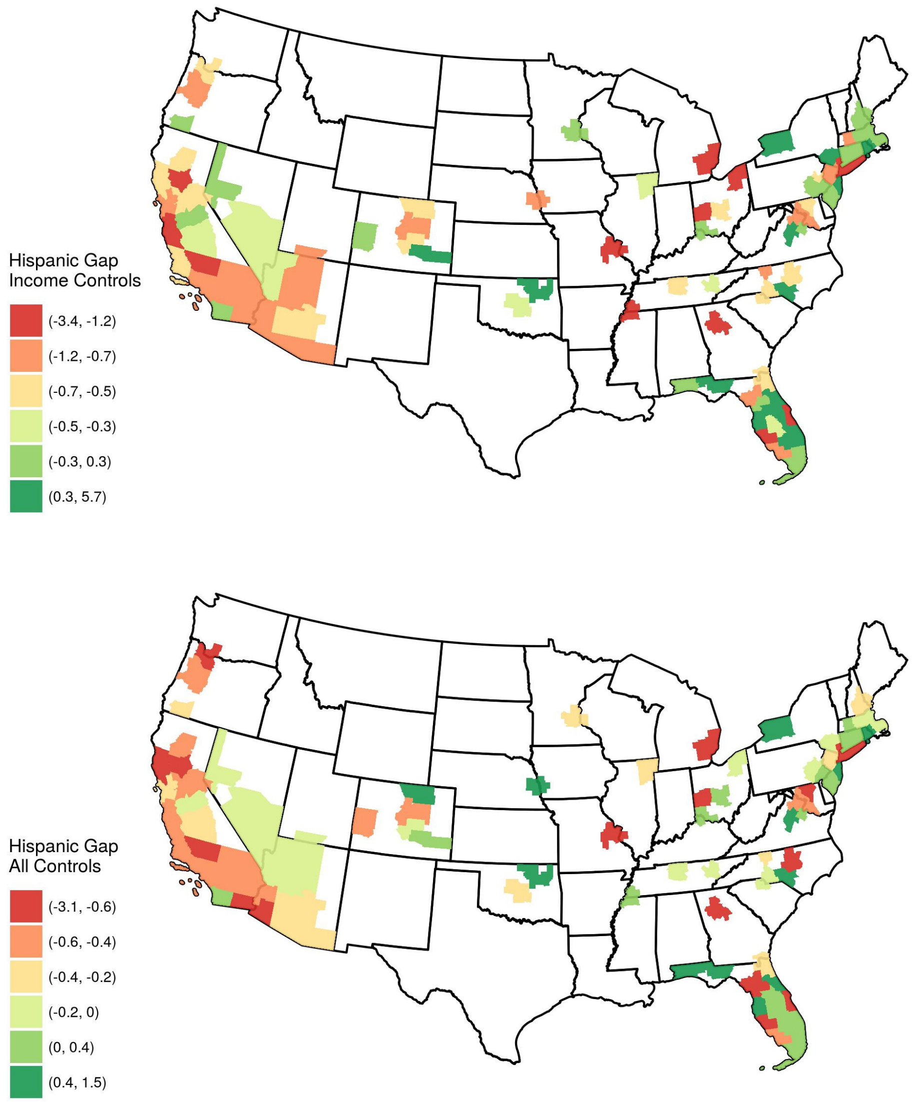


Figure 2.15: Map of the Hipsanic-white appreciation gap by commuting zone. Areas in red have particularly problematic gaps, while areas in green have much less problematic—or even reversed—gaps. The top panel measures gaps relative to predictions based only on income and date, while the bottom panel measures gaps relative to predictions based on all variables discussed in Section 2.4.

map considers the gap relative to predictions based only on income, while the lower map considers the gap relative to predictions based on the full complement of controls in Section 2.4. The black-white gap appears particularly bad through the South and Midwest, with much smaller gaps on the two coasts. This is consistent with intuition regarding racial segregation, but less consistent with the findings on interracial gaps in income mobility in Chetty et al. (2018), which find poor mobility for blacks and in a way that is much more geographically uniform or sporadic (see their Figure IX and Online Appendix Figures X and XI). The Hispanic-white gap appears much more sporadic. In both Figures, it is clear that the two measures of the appreciation gap are very similar, so from this point forward I consider only the one that controls flexibly for all observables—the one in the bottom map of each Figure. Figure 2.16 graphs the

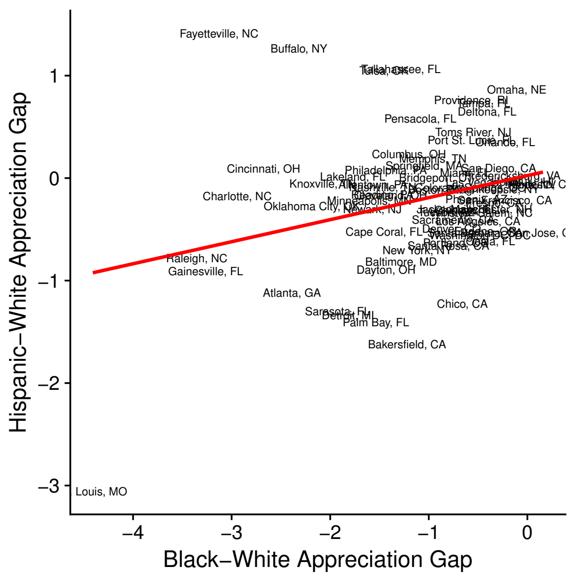


Figure 2.16: Relationship between the black-white appreciation gap and the Hispanic-white appreciation gap across commuting zones.

relationship between the black-white and Hispanic-white gaps, and shows a positive but weak correlation.

One might ask whether these gaps are correlated with demographic and economic attributes of the cities. Figures 2.17 and 2.18 help answer that question by plotting

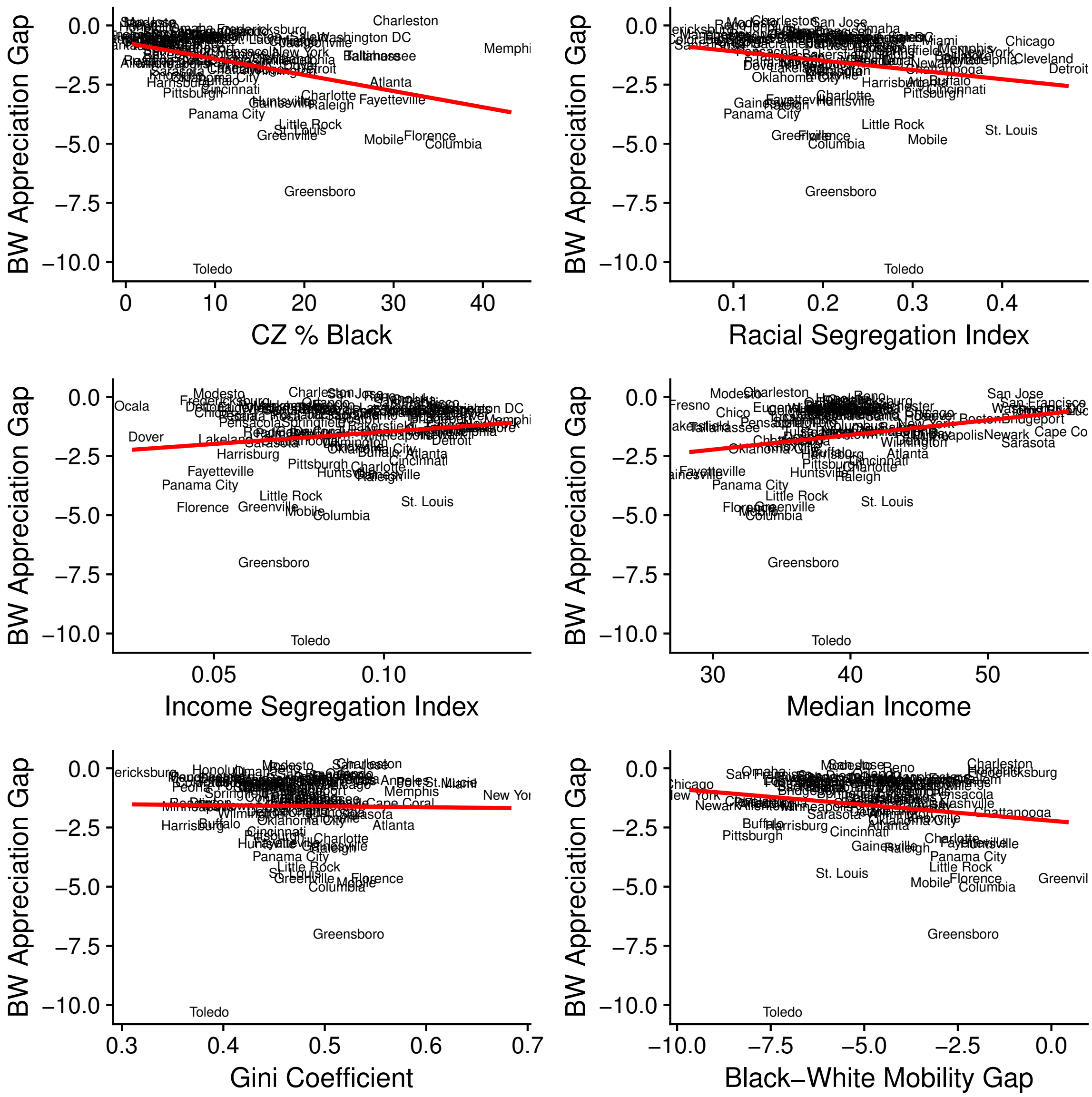


Figure 2.17: Relationship between the black-white appreciation gap and many salient economic attributes of commuting zones. All horizontal axis variables are taken from online tables associated with Chetty et al. (2014) and Chetty et al. (2018). Racial segregation is a Thiel index, while income segregation is a Rank-Order index. The black-white mobility gap is taken as the black-white gap in average offspring income percentile, averaged over 4 groups: males born to 75th percentile parents, males born to 25th percentile parents, females born to 75th percentile parents, females born to 25th percentile parents.

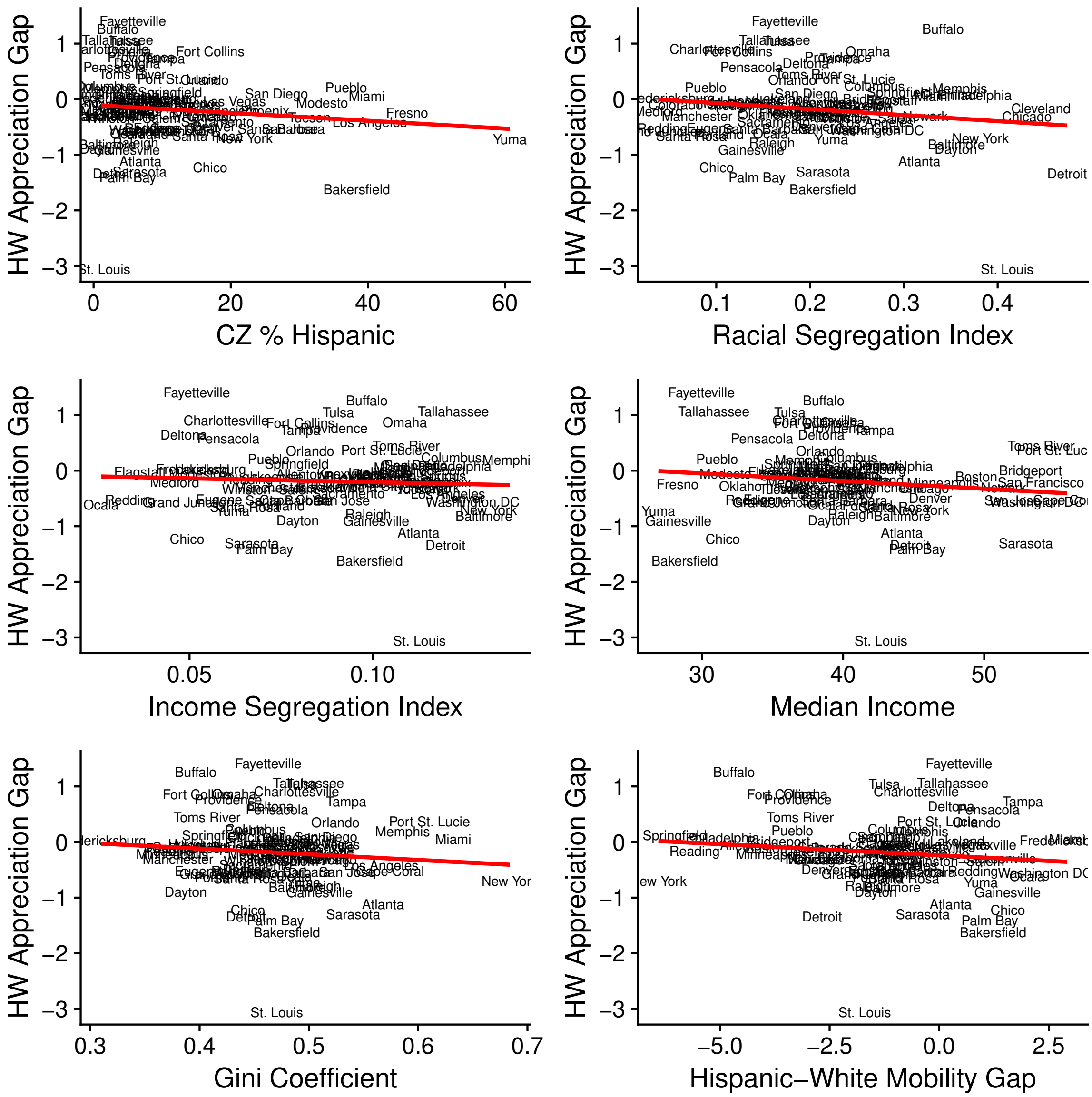


Figure 2.18: Relationship between the Hispanic-white appreciation gap and many salient economic attributes of commuting zones. All horizontal axis variables are taken from online tables associated with Chetty et al. (2014) and Chetty et al. (2018). Racial segregation is a Thiel index, while income segregation is a Rank-Order index. The Hispanic-white mobility gap is taken as the Hispanic-white gap in average offspring income percentile, averaged over 4 groups: males born to 75th percentile parents, males born to 25th percentile parents, females born to 75th percentile parents, females born to 25th percentile parents.

the black-white and Hispanic-white gaps, respectively, against many such attributes. Overall, it appears that the appreciation gaps are unrelated to most of these attributes; the only exceptions that stand out are that the black-white appreciation gap appears most extreme in cities with more black residents and cities that are more racially segregated.

In summary, while the degree of the interracial appreciation gap varies strongly across the country, it does not appear to do so in a way that is systematically correlated with other important variables that might be intuitively tied to racial inequality.

2.6 Conclusion

In the foregoing sections, I have demonstrated a large gap of around 2.5 percentage points per year in the home price appreciations experienced by black and white homeowners, even after flexibly accounting for income. Hispanics earned returns around 1.5 percentage points lower than whites. I then showed that, despite large differences in the types of homes and neighborhoods selected by homeowners of different races but the same income, these differences cannot explain the entire gap in appreciation. Flexibly accounting for all of these characteristics still leaves a gap of around 1 percentage point per year. Finally, I investigated whether this gap is larger in some portions of the country and found that to be the case. However, the severity of the gap is mostly uncorrelated with other economic characteristics of the city such as income inequality and the interracial mobility gap measured by Chetty et al. (2018).

As extensive as this analysis was, three major areas remain for further investigation. First, Figure 2.3 shows that data coverage is remarkably uneven across the country. While most major population centers are covered well enough for study, Texas, New Orleans, and Seattle stand out as places with severely lacking data. Other, commercially available, superior data sources exist, and the analysis could be repeated. Texas, which contains the two largest metro areas in the South, would be a particularly important area for study.

Second, one could look further into possible explanations for this effect. Possibilities include physical differences in the homes selected by different races, differential investment during the homeownership spell, amenities of the neighborhoods selected, and even long standing consequences of redlining. Finally, and most mechanically, homeowners of underrepresented race or ethnic groups may simply be overpaying for homes they purchase or underpricing homes they sell. To answer this question, one must build a strongly accurate prediction for home price as a function of location, date, and physical characteristics, and test for whether deviations from that prediction are correlated with race.

Third, documentation and explanation of this phenomenon are important in drawing attention to it. However, since this appreciation gap clearly contributes substantially to interracial economic disparity, researchers should investigate whether and which policies can mitigate it. This might be approached via study of repeat sales transactions, linked to demographic variables about the homeowner, around periods of policy change or natural experiments in local areas.

Chapter 3

The Effect of State Income Taxes on Home Values: Evidence from a Border Pair Study

3.1 Introduction

State and local governments account for about 40% of all tax collections in the United States (Williams, 2012), but federal taxes command most of the attention in academic literature. In this paper, I investigate the effect of state income taxes on home prices. Empirically, I provide suggestive evidence that the elasticity of home prices with respect to state income taxes is large. Ultimately, however, the evidence is inconclusive; standard errors are large, and different specifications lead to different conclusions. Nonetheless, using benchmark point estimates of the elasticity, I argue that ignoring general equilibrium effects on other prices and quantities when evaluating government policy can lead to large errors in the calculation of the marginal value of public funds (MVPF) associated with a policy. In particular, further work should be performed on this particular elasticity to obtain a more precise estimate.

In the main empirical sections of the paper, I employ four variants of a difference-in-differences (DD) strategy: First, I consider a standard DD, regressing log home prices on log net-of-tax rates with ZIP code and time fixed effects. Second, I use the border ZIP code pair approach used by Dube, Lester and Reich (2010), in which time fixed effects are replaced by time-pair fixed effects, essentially controlling for home prices in treated ZIP code A using home prices in untreated, adjacent ZIP code B. Third, to address the issue of causation and the validity of the parallel trends assumption, I use a distributed leads/lags model, regressing log home price on 3 years' worth of leads and lags of monthly changes in the tax differential (as well as ZIP code and time-pair fixed effects). Finally, since most tax changes are small (a standard deviation of only 0.15% over the sample period), I single out several case studies that involve substantial tax changes and relatively complete home price data. In each case, I regress log home price differential on the interaction of time-fixed effects and a dummy that is 1 (plus ZIP code and time-pair fixed effects). Though many variants yield large point estimates, some do not, and standard errors are typically large. Thus, I find suggestive, but ultimately inconclusive, evidence of a relationship.

The intuition behind the conceptual insight—that the MVPF must include the general equilibrium (GE) effects of other prices and quantities—is as follows. When state income taxes rise, the state becomes a less attractive place in which to reside. To ensure housing markets clear, the price of homes within the state falls. Thus, in addition to the usual behavioral response to the increased labor tax, there is an additional general equilibrium effect on the government budget, if the government collects property taxes as well. Though this GE effect appears in both the numerator and denominator—it reduces government revenue, but through a direct transfer of those resources to individuals—it still increases the MVPF, because while the usual behavioral response to the increased labor tax is unaffected, the policy is "smaller." That is, it both raises less revenue, and reduces individual utilities by less, but revenue leakage is the same. Said another way, accounting for GE effects on home prices does not change the excess burden, in dollar terms, imposed by state income taxes but does reduce the amount of revenue collected, making the policy less attractive.

The rest of the paper is organized as follows: In the rest of this section, I discuss the relevant literature and explain more fully how this paper contributes to and advances it. Section 3.2 describes more fully the motivation for the paper, and the theoretical framework behind the empirical specifications. Section 3.3 describes the two data sources employed by the paper, and presents summary statistics. Section 3.4.1 presents the specification and results for the standard border-pair difference-in-differences estimator. Section 3.4.2 lays out the analysis of the dynamic effect of state income taxes on home prices and presents the results. Section 3.4.3 carefully describes the situations surrounding the two case studies and the outcomes. Section 3.5 discusses the implications of these results, and ties them back to the original motivation for the paper. Finally, Section 3.6 concludes.

Related Literature At the most superficial level, this paper might be seen as the latest in a long line of empirical papers that implement a key theoretical argument: Chetty (2009) argues that one can often use clever theoretical arguments to show the welfare relevance of simple sufficient statistics, such as elasticities. In the article, Chetty highlights past applications of this idea in numerous areas, from tax policy to social insurance to labor economics. Hendren (2016) takes the argument one step further by noting that, in most cases, the welfare relevant statistic is simply the effect of behavioral responses to government policies on the government budget, or what Hendren calls the "fiscal externality." This observation is particularly relevant for my investigation. The obvious fiscal externality caused by higher income taxes is a reduction in taxable income that results from decreased incentives. But when the income taxes are local, the logic of compensating differentials ensures that local utility is pinned down by surrounding utility; thus, local home prices will fall, thereby decreasing property tax collections. Assuming property tax millage rates[1] do not change in a coordinated fashion, the elasticity of home prices with respect to local income taxes is equivalent to the elasticity of property tax collections with respect to local income taxes and may be an important part of Hendren's "policy elasticity." It turns out that this effect actually shows up not as a fiscal externality, but instead as a reduction in the tax base, while holding the fiscal externality constant; thus, the fiscal externality erases a greater portion of the revenue raised by the policy.

This paper also contributes to at least three major strands of empirical literature, the first of which should be seen as an application of the above theory: a series of papers that estimate elasticities of certain real economic activity with respect to a given tax rate. A canonical example is Feldstein (1995), which estimates the elasticity of taxable income (found to be the welfare-relevant elasticity) with respect to the (marginal) net-of-tax rate. Feldstein finds a large elasticity, possibly exceeding 2, by using a small panel of taxpayers, thereby controlling for individual effects. Saez (2010) uses information contained in a single cross-section to infer this elasticity, arguing that the amount of bunching at kink points in the tax schedule maps to this elasticity. Many other similar studies, each exploiting different natural experiments, exist, including Eissa (1995), Fortin, Lemieux and Frechette (1994), and Goolsbee (2000). Saez, Slemrod and Giertz (2012) have also contributed a thorough review of this literature. This paper looks specifically at the elasticity of home prices with respect to state income tax rates.

The second investigates the capitalization of economic conditions and taxes into

[1]the percentage of assessed property value that is levied in property taxes each year

asset prices. Cutler (1988) applies this idea to the stock market response to the Tax Reform Act of 1986—the same Act whose effects Feldstein (1995) exploits as a natural experiment. In addition to this Act lowering top marginal personal income tax rates, it alters various aspects of corporate tax and dividend tax policy. Cutler argues that while mechanical changes in cash flows as a result of the policy changes should be correlated with excess returns, so should other aspects of the company's balance sheet. For example, repealing the investment tax credit makes new investment less attractive, thereby driving up the price of existing capital in general equilibrium; companies with a substantial stock of such capital should benefit, and the data bear this out. In the present situation, it is the home prices in the state enjoying lower income taxes that should rise in general equilibrium. Other papers demonstrating this capitalization include Friedman (2009) and Linden and Rockoff (2008), with the latter especially relevant because it focuses on home prices' response to the presence of sex offenders.

The third focuses specifically on the issue of state and local income taxes, though not through the traditional lens of welfare-relevant elasticities. One important paper in this thread is Feldstein and Wrobel (1998), which argues that states cannot redistribute income by showing that wages adjust upward (and, assuming the labor demand curve doesn't shift, the quantity of people employed adjusts downward) to compensate for higher taxes. The identification, however, comes only from instrumenting for individual tax liabilities *given* state tax rates; the state tax rates are taken as exogenous. Young and Varner (2011) directly estimate, using microdata, the tendency of the rich to emigrate to evade a high tax rate. It finds a small propensity in the case of the New Jersey "millionaire" tax. However, the approach fails to account for the adjustment of home prices (especially home prices aimed at the wealthy) in general equilibrium, which may absorb most of the shock. The present paper takes up precisely this adjustment.

3.2 Theoretical Framework

3.2.1 Structural Model

Before starting on the empirical estimation or delving into the welfare motivation for this study, I here provide a foundation for the specifications that I use in the empirical sections. This will be helpful in fixing ideas and notation before the welfare section to follow. The model, however, should be taken as primarily evocative and pedagogical rather than empirically precise.

Consider a set of identical individuals who each have the option of living in state $j = 1$ or $j = 2$. State j has a linear tax rate τ_j^L on labor and τ_j^H on property. Each individual earns wage w regardless of which state he chooses to live in. Once choosing a state in which to live, he selects an amount of housing H to consume at pre-tax price h_j per unit; other consumption C has price 1. He also decides on how much work effort to supply, L. He faces the budget constraint

$$C + Hh_j(1 + \tau_j^H) = w(1 - \tau^L)L.$$

One state may be more desirable than the other, due either to pure preferences or local services; the desirability of state j is captured by u_j.

Conditional on choosing state j, the individual maximizes his utility

$$U_j = u_j \left\{ C^a H^b - \frac{\theta}{1+\gamma} L^{1+\gamma} \right\},$$

where $a > 0, b > 0, \gamma > 0, \theta > 0, a + b \leq 1$, subject to the budget constraint above. Solving the individual's maximization problem yields the following:

Proposition 3.2.1 *Individuals chose the following policy functions:*

$$L(w(1-\tau_j^L), h_j(1+\tau_j^H)) = \tilde{L} \cdot (h_j(1+\tau_j^H))^{\frac{-b}{\gamma+1-a-b}} (w(1-\tau_j^L))^{\frac{a+b}{\gamma+1-a-b}}$$
$$C(w(1-\tau_j^L), h_j(1+\tau_j^H)) = \tilde{C} \cdot (h_j(1+\tau_j^H))^{\frac{-b}{\gamma+1-a-b}} (w(1-\tau_j^L))^{\frac{\gamma+1}{\gamma+1-a-b}}$$
$$H(w(1-\tau_j^L), h_j(1+\tau_j^H)) = \tilde{H} \cdot (h_j(1+\tau_j^H))^{\frac{-(\gamma+1-a)}{\gamma+1-a-b}} (w(1-\tau_j^L))^{\frac{\gamma+1}{\gamma+1-a-b}}$$

and obtain the following indirect utility:

$$V(w(1-\tau_j^L), h_j(1+\tau_j^H), u_j) = u_j \tilde{V} \cdot (h_j(1+\tau_j^H))^x (w(1-\tau_j^L))^y$$

where $x = \frac{-b(\gamma+1)}{\gamma+1-a-b}$, $y = \frac{(a+b)(\gamma+1)}{\gamma+1-a-b}$, *and* $\tilde{L}, \tilde{C}, \tilde{H}$, *and* $\tilde{V}$ *are constants, across space and policy changes.*

Proof. See Appendix C.1. ■

If both states 1 and 2 have constrained housing supplies, then for the housing market to clear, individuals must be indifferent across space, yielding

$$\begin{aligned} u_1\tilde{V} \cdot (h_1(1+\tau_1^H))^x (w(1-\tau_1^L))^y &= u_2\tilde{V} \cdot (h_2(1+\tau_2^H))^x (w(1-\tau_2^L))^y \\ \ln h_1 - \ln h_2 &= -\frac{1}{x}(\ln u_1 - \ln u_2) - (\ln(1+\tau_1^H) - \ln(1+\tau_2^H)) \\ &\quad - \frac{y}{x}(\ln(1-\tau_1^L) - \ln(1-\tau_2^L)) \end{aligned}$$

More generally, if there are S states in which to live, all with constrained housing supply, and this market clears at many points in time t, then we must have

$$\begin{aligned} \ln h_{jt} =& \ln h_{1t} + \frac{y}{x}\ln(1-\tau_{1t}^L) - \frac{y}{x}\ln(1-\tau_{jt}^L) \\ & - \frac{1}{x}(\ln u_{jt} - \ln u_{1t}) - (\ln(1+\tau_{jt}^H) - \ln(1+\tau_{1t}^H)) \\ =& \psi_t + \eta \ln(1-\tau_{jt}^L) + \tilde{u}_{jt} \end{aligned}$$

where

$$\psi_t \equiv \ln h_{1t} + \frac{y}{x}\ln(1-\tau^L_{1t})$$
$$\eta \equiv -\frac{y}{x}$$
$$\tilde{u}_{jt} \equiv -\frac{1}{x}(\ln u_{jt} - \ln u_{1t}) - (\ln(1+\tau^H_{jt}) - \ln(1+\tau^H_{1t})).$$

$\tilde{u}_{jt}$ can be further broken down into an average ϕ_j and time-varying component μ_{jt}, which yields the standard difference-in-differences equation

$$\ln h_{jt} = \phi_j + \psi_t + \eta \ln(1-\tau^L_{jt}) + \mu_{jt} \tag{3.1}$$

Estimating this equation—especially η—is the purpose of the empirical sections of this paper.

3.2.2 Welfare Motivation

As mentioned in the Introduction, Hendren (2016) defines the marginal value of public funds (MVPF) associated with a particular policy affecting a homogeneous group of people as the ratio of their dollar-equivalent reduction in utility per dollar of revenue collected by the government for "small" versions of the policy. He argues that the MVPF for a pure tax policy can, in the absence of general equilibrium effects, be written as

$$MVPF = \frac{\text{Marginal Mechanical Revenue}}{\text{Marginal Actual Revenue}} = \frac{1}{1-FE}$$

where FE is the "fiscal externality"—the revenue lost by the government due to behavioral response to the policy. This is derived through use of the envelope theorem. Optimal policy can be achieved by setting the social MVPF—the MVPF weighted by subjective social marginal utilities of income—equal along all possible policy paths. In the Appendix, he quickly notes that general equilibrium effects on prices should be included if they exist. Here, I take up the subject of what that looks like in this particular case.

Consider increasing the labor tax rate slightly in a particular state, by ϵ. Recalling from the previous subsection that, in the short run, no one moves as a result of this policy, this has two effects on an individual's utility, after applying the envelope theorem: its mechanical tax effect, and its general equilibrium effect on housing prices in his chosen state.[2] The dollar-equivalent utility reduction from the first effect is simply $\epsilon w L(w(1-\tau_j^L), h_j(1+\tau_j^H))$. The dollar-equivalent utility increase from the second effect is $\frac{\epsilon}{1-\tau_j^L}\eta h_{jt}(1+\tau_j^H)H(w(1-\tau_j^L), h_j(1+\tau_j^H))$. However, this also has an effect on whoever owns the property, and leases it to the individual; it decreases his revenue by $\frac{\epsilon}{1-\tau_j^L}\eta h_{jt}H(w(1-\tau_j^L), h_j(1+\tau_j^H))$.[3]

Moving on to the government budget, the increase in tax has three separate effects: the mechanical revenue effect, the mechanical effect of the drop in home prices, and the behavioral response to these price changes. The behavioral response can be further decomposed into four effects: the labor supply and housing consumption effects of the changes in net-of-tax wage and home prices. The mechanical effects are $\epsilon w L(w(1-\tau_j^L), h_j(1+\tau_j^H))$ and $-\frac{\epsilon}{1-\tau_j^L}\eta h_{jt}\tau_j^H H(w(1-\tau_j^L), h_j(1+\tau_j^H))$, respectively.

---

[2]In this exercise, I assume that any change in the *difference* in log housing prices occurs in the state adjusting its policy. Presumably there are "third party" states as well, which pin down log housing prices in the second state.

[3]I assume the landlord has no behavioral response, and is merely a large, risk-neutral corporation that rebates his profits lump-sum to individuals.

The behavioral responses' effects on the budget are as follows:

$$\begin{aligned}
\text{Labor Supply to Wage} \quad & -\frac{\epsilon}{1-\tau_j^L} e_L^w w \tau_j^L L(w(1-\tau_j^L), h_j(1+\tau_j^H)) \\
\text{Labor Supply to Housing Prices} \quad & -\frac{\epsilon}{1-\tau_j^L} \eta e_L^h w \tau_j^L L(w(1-\tau_j^L), h_j(1+\tau_j^H)) \\
\text{Housing Consumption to Wage} \quad & -\frac{\epsilon}{1-\tau_j^L} e_H^w h_j \tau_j^H H(w(1-\tau_j^L), h_j(1+\tau_j^H)) \\
\text{Housing Consumption to Prices} \quad & -\frac{\epsilon}{1-\tau_j^L} \eta e_H^h h_j \tau_j^H H(w(1-\tau_j^L), h_j(1+\tau_j^H))
\end{aligned}$$

where e_L^w is the (direct) elasticity of labor supply with respect to net-of-tax wage, and other es are defined similarly.

Notice that we can replace $w\tau_j^L L_i$ with R_i^L, the revenue from the labor tax collected from individual i and $h_j \tau_j^H H_i$ with R_i^H, the revenue from the property tax collected from individual i. Thus, we have arrived at the following proposition:

Proposition 3.2.2 *The MVPF of raising the state income tax is given by*

$$MVPF = \frac{\frac{1}{\tau^L} R^L - \frac{\eta}{1-\tau^L} R^H}{\frac{1}{\tau^L} R^L - \frac{1}{1-\tau^L}\left[\eta R^H + \left(\underbrace{e_L^w + \eta e_L^H}_{\textit{Total response}}\right) R^L + \left(\underbrace{e_H^w + \eta e_H^h}_{\textit{Total response}}\right) R^H\right]} \tag{3.2}$$

If housing is not considered, except that the total labor supply behavioral response $(e_L^w + \eta e_L^h)$ is correctly measured, then the MVPF would simply reduce to

$$MVPF = \frac{1}{1 - \frac{\tau^L}{1-\tau^L}(e_L^w + \eta e_L^h)}$$

which is the standard form.

After empirically estimating η in the ensuing sections, I will return to this MVPF calculation in Section 3.5.

3.3 Data

This paper primarily employs data from two sources: Home price data comes from Zillow Research (www.zillow.com/research/data/), while tax data comes from published output of the NBER TAXSIM model (users.nber.org/∼taxsim/). These data are supplemented with population and geography data from the Census.

3.3.1 Population and Geography

Data on the 2010 population and 2016 location of all 33,144 U.S. ZIP Code Tabulation Areas (ZCTAs)[4] were gathered from NHGIS (Manson et al., 2017) and Census TIGER/Line®Shapefiles,[5] repectively. The adjacency matrix was then computed using the R package "sf." I define a *border pair* as a pair of ZIP codes that border each other but are not members of the same state.[6] I then treat these ZCTAs as ZIP codes in a naive fashion.

3.3.2 Home Prices

Zillow Research publicly publishes various home price indices at the state, metro, county, city, ZIP code, and neighborhood levels. These home price indices include various percentiles of home prices, condo prices, single-family home prices, value per square foot, and prices for various subsets of single family homes. In this paper, I consider

[4] It is worth noting that ZIP codes are not actual places or geographical polygons but, rather, lists of addresses. In general, these addresses are geographically clustered, but some of them are a single point (a P.O. box, for instance) and some are not contiguous. The Census, instead of the actual ZIP codes, uses ZCTAs—ZIP code tabulation areas—which are geographical polygons

[5] via R package "tigris"

[6] A handful of ZIP codes do cross state lines. I use the Census relationship files to find what percentage of the ZCTA's population resides in each state, and then drop any ZCTAs for which no state contains 90% of the population.

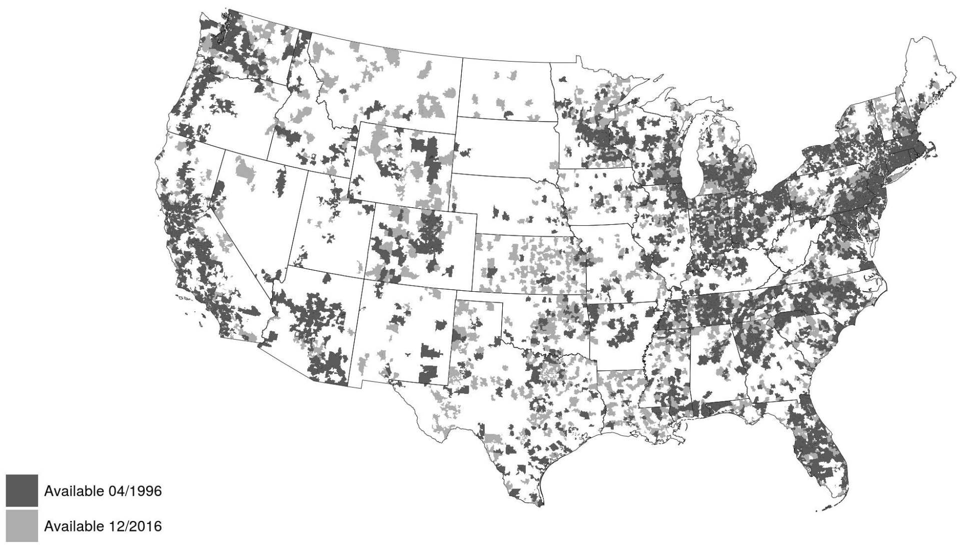


Figure 3.1: Availability of Zillow data. Dark areas have median home price data since April 1996, while light areas do not have median home price data then but do by December 2016. The white areas have no home price data.

only the median home prices and the median value per square foot, the latter of which is more closely tied to h_{jt} in Equation 3.1. These data are monthly from April 1996 through 2017, though not every ZIP code has data in any given month.

Figure 3.1 shows the availability of median home price data by ZIP code. While much of the geographical area of the country is not covered by the dataset, there is substantial coverage of the most populated areas of the country. Specifically, Figure 3.2 shows that the vast majority of the country's population is covered by these data as far back as 1997, with increasing coverage since then. Over 50% of that population living in a ZIP code on a state border is covered for the entire sample period, and over 65% by 2002.

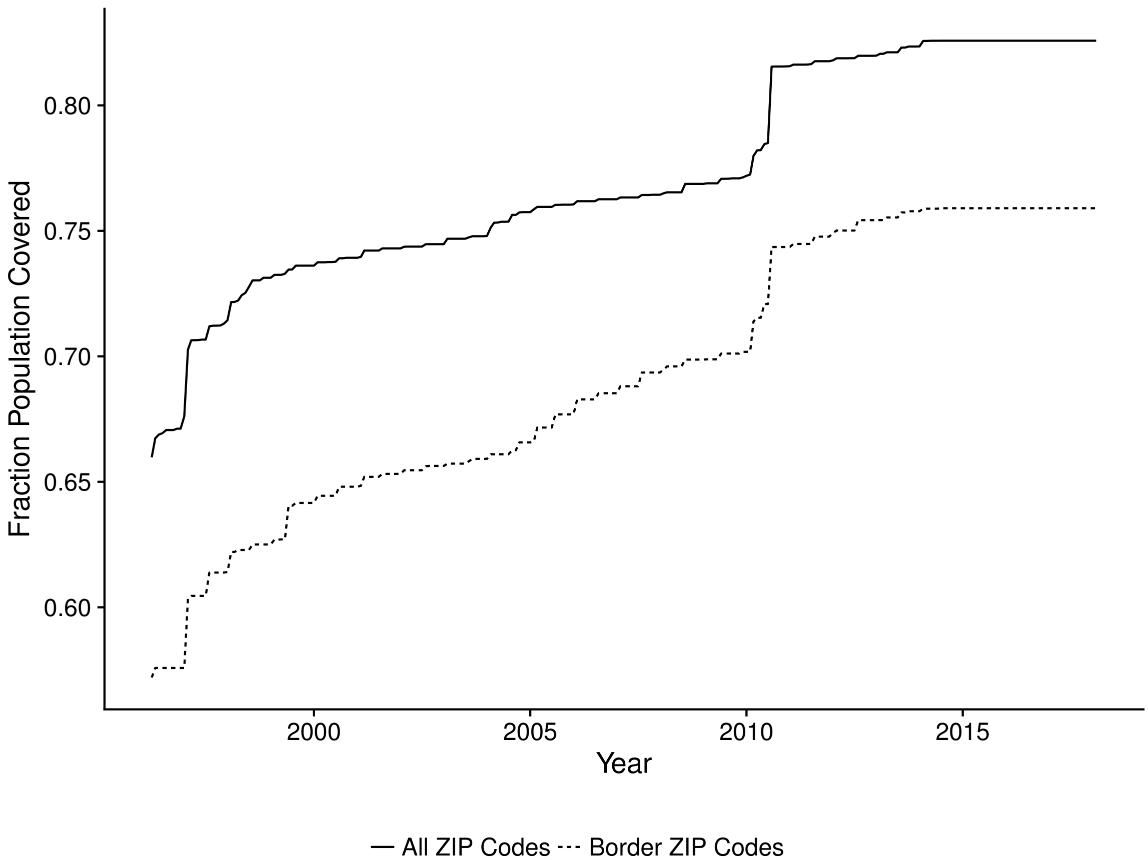


Figure 3.2: Population covered by Zillow data. In each month, the 2010 populations of the ZIP codes which Zillow covers in that month are summed and divided by the sum of the 2010 populations of *all* ZIP codes to arrive at the point on the graph. Thus, the upward trend in the graph is due only to the increasing quality of data, and not to population growth.

Table 3.1: Home price summary statistics

	1997			2017		
Variable	N	Mean	SD	N	Mean	SD
All ZIP Codes						
Median Price	11358	119.4	69.7	15362	245.2	240.1
Median Price/Sq. Ft.	11380	74.2	33.3	14771	156.3	141.3
Border ZIP Codes						
Median Price	997	106.6	61.3	1463	190.8	163.7
Median Price/Sq. Ft.	995	66.2	27.0	1381	120.2	83.7

Statistics for the entire sample of available ZIP codes, as well as for the subsample of ZIP codes that are members of a border pair, are reported. Median prices are in thousands of dollars, while prices per square foot are in dollars.

Summary statistics for the Zillow data can be found in Table 3.1.[7] As in Dube, Lester and Reich (2010), I present summary statistics from the home price data using two samples: the entire sample of ZIP codes for which data is available, and the sample that includes only those ZIP codes that are part of a border pair. In both cases, one can see that there is similar coverage in the "median value" and "median value per square foot" variables. The samples are large, with even the smallest sample (1997 data on ZIP codes in a border pair) containing 995 ZIP codes. Additionally, the data appears to be fairly representative, with the values at least qualitatively in line with national home prices.

3.3.3 Taxes

NBER's TAXSIM model (Feenberg and Coutts, 1993) is a piece of software that calculates tax liabilities for tax units, given the various data that would be collected on

[7] I have chosen the beginning and end points merely to be informative; this nationwide study is not formally a "difference-in-differences" study, and therefore there is no formal "before and after."

a tax return. In addition to being able to calculate such liability for any hypothetical tax unit a researcher might wish to study, various liabilities of interest have already been calculated and published on the Web for every state and year from 1977 through 2016.

The theoretically relevant tax liability for a household deciding its state of residence is the *total*, not marginal, tax that would be owed in the states up for consideration. For robustness, I consider three measures of this, all of which are provided in the published TAXSIM tables. Before I detail these three measures, however, it is worth noting that whether I use the state tax rate or the total (state plus federal) tax rate is irrelevant. Since state taxes are deductible on federal tax returns but not vice-versa,[8] the net-of-tax total rate is $1 - \tau_T = (1 - \tau_f) \cdot (1 - \tau_s)$. Taking logs (which I do in all of my specifications, so that regression coefficients can be interpreted as elasticities), we have $\log(1 - \tau_T) = \log(1 - \tau_f) + \log(1 - \tau_s)$. Since all specifications involve a time fixed effect, $\log(1 - \tau_f)$ can be dropped.

For the first two measures, I consider the state tax owed by a typical family with two adults and two children at nominal incomes of \$50,000 and \$100,000. The dataset makes assumptions about the nature of such families' income (what percentage is wages, the extent of their deductions, etc.). Since I've chosen to use the taxes owed at a constant nominal income, changes in the taxes owed are due only to changes in tax law and not to inflation.[9] I've chosen a 4-person family because families comprise the primary market for owner-occupied housing. These incomes I've chosen represent the 43rd and 72nd percentiles, respectively, of the 2016 U.S. income distribution (U.S.

[8]In a few states, federal tax is deductible on the state tax return, but this is accounted for in the data.

[9]A few states do index their brackets to inflation, but they are in the minority. Additionally, inflation has been low throughout the sample period, so the indexing of the brackets in these states will be relatively unimportant.

Table 3.2: Tax summary statistics.

	1997			2015		
Tax Rate	N	Mean	SD	N	Mean	SD
Levels						
Family of 4 with $50k Income	51	3	1.8	51	2.2	1.5
Family of 4 with $100k Income	51	3.8	2.1	51	3.2	1.8
Top Marginal Rate	51	5.4	3.1	51	5.2	3.2
Year-Over-Year Changes						
Family of 4 with $50k Income	1020	-0.05	0.15			
Family of 4 with $100k Income	1020	-0.03	0.15			
Top Marginal Rate	969	-0.01	0.32			

Statistics on levels in 1997 and 2015 are presented, followed by statistics on the pooled year-over-year changes for all years since 1997. The "family of 4" measures are average tax rates, while the top marginal rate is, of course, a marginal rate. They are measured in percentage points.

Census Bureau, 2017), and therefore are probably typical of middle class home buyers. Some counties or ZIP codes are marketed to significantly wealthier households. Thus, I've also used the top marginal rate on wages as a proxy for average taxes faced by these wealthier households. These top marginal rates are calculated only through 2015.

Table 3.2 presents summary statistics on the average tax rates, as well as their year-over-year changes. The levels clearly have lots of heterogeneity (standard deviations around 2 percentage points), while the year-over-year changes are more homogeneous (standard deviations around 0.15 percentage points). It is this homogeneity that will lead to some of the methodological issues this paper presents. Note, however, that the standard deviation of the year-over-year changes is still large relative to the mean.

Table 3.3: Simple difference-in-differences estimates

Tax Rate	Median Home Price			Per Sq. Ft. Home Price		
Family of 4 w/ $50k Income	5.668 (5.403)			6.993** (3.279)		
Family of 4 w/ $100k Income		7.373* (3.811)			8.078** (3.666)	
Top Marginal			−3.085*** (0.831)			−3.018*** (0.795)
Observations	3,418,861	3,418,861	3,234,697	3,360,430	3,360,430	3,183,370
R^2	0.952	0.953	0.952	0.944	0.944	0.943
Adjusted R^2	0.952	0.952	0.952	0.943	0.944	0.943
Note:				*p<0.1; **p<0.05; ***p<0.01		

Estimates of (3.3) using the full ZIP code sample. All dependent and independent variables are in terms of logs. The unit of observation is ZIP code-month. Coefficients should be interpreted as elasticities. All independent variables are average state net-of-tax rates on the population listed. All models include ZIP code and time fixed effects. Standard errors, clustered at the state level following Bertrand, Duflo and Mullainathan (2004), are in parentheses.

3.4 Empirical Results

3.4.1 Static Results

The empirical goal of this paper is to estimate (3.1), which I reproduce here for convenience:

$$\ln h_{jt} = \phi_j + \psi_t + \eta \ln(1 - \tau_{jt}^L) + \mu_{jt}. \tag{3.3}$$

Straightforward OLS estimates of this specification can be found in Table 3.3. Clearly, the relationship is economically significant, though the standard errors are large enough to make meaningful inference difficult, except in the case of top marginal rates, which have the opposite of expected sign. We will see that large standard errors remain a common problem throughout. More important, these estimates are valid only if μ_{jt} is uncorrelated with the local tax rate, conditional on the ZIP code and time fixed effects. Recall from Section 3.2 that μ_{jt} encapsulates two economic forces: relative desirability—aside from taxes—of the two regions and the difference in property taxes. Property tax rates are highly local, whereas income tax rates vary mostly at the state level. While these rates may be statically anticorrelated—foregone labor tax revenue must be made up for by property tax revenue—changes in the two (i.e., accounting for the ZIP code fixed effect) are far less likely to be; unfortunately, gathering data on local property tax rates is beyond the scope of this paper, but should be investigated in further research.

Most concerning is possible correlation between relative net-of-tax rates and relative desirability of the two ZIP codes, which is likely to be present. ZIP codes that become more attractive for reasons having nothing to do with taxes might simultaneously see a drop in tax rates, as the government can pay its expenses with a lower rate during a local boom. Thus, estimates of (3.3) incorrectly interpret this correlation as a causal effect of tax rates on home prices.

To combat this, I propose border pair, difference-in-differences techniques. ZIP codes that are neighboring, but in different states, should experience similar shocks to their desirabilities. This can be accounted for with a pair-time fixed effect,[10] which controls for general conditions in the area, including desirability. I employ the classic border pair specification proposed in Dube, Lester and Reich (2010):

$$\ln h_{jpt} = \phi_j + \psi_{pt} + \eta \log(1 - \tau_{jt}) + \mu_{jpt}, \tag{3.4}$$

where p indexes the border pair and ψ_{pt} is a pair-time fixed effect.

[10]Formally, assume that μ_{jt} can be decomposed into a pair-time fixed effect ψ_{pt}, which captures changes in the general area's desirability, and μ_{jpt}, the residual. Then one can subsume the original ψ_t into this new ψ_{pt}.

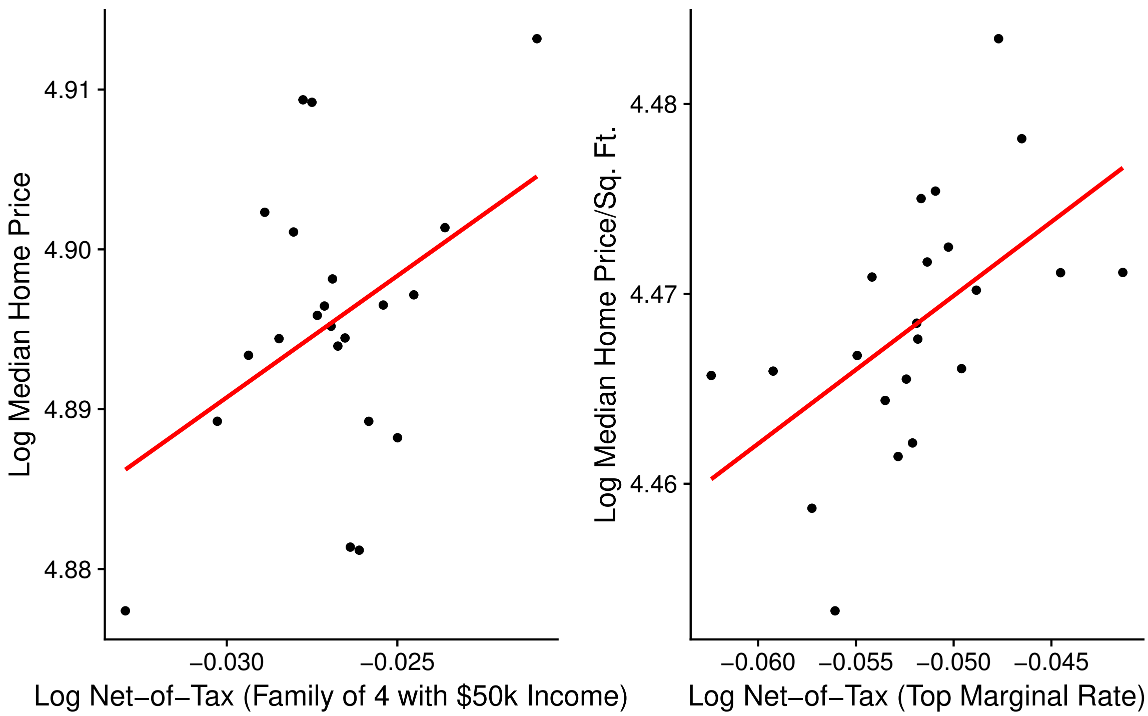


Figure 3.3: Per-square-foot home prices vs. net-of-tax rates for two different definitions of tax rate. Both variables have had logs taken, meaning the slope should be interpreted as an elasticity. Both variables have been residualized against ZIP code and pair-time fixed effects

Figure 3.3 shows binscatters[11] of log per-square-foot price versus net-of-tax rates, controlling for ZIP code and pair-time fixed effects. The relationship appears to be strong. Table 3.4 presents the regression results. The residuals are subject to various types of correlation across observations as noted by Dube, Lester and Reich (2010). First, home prices are measured monthly and at the ZIP code level, while taxes are measured annually and at the state level. This implies errors may be correlated for two observations within the same state-year. In addition, pairs along a border segment (pair of neighboring states) may mechanically have correlated errors, since the same ZIP

[11]Observations are grouped into 40 equally sized bins along the horizontal axis, and the mean of both the x-axis variable and the y-axis variable within each bin is plotted as a single point. This is helpful in absorbing much of the noise of the data, preserving the overall pattern.

Table 3.4: Estimates with pair-time fixed effects

Tax Rate	Median Home Price			Per Sq. Ft. Home Price		
Family of 4 with $50k Income	1.444 (1.246)			2.327** (1.161)		
Family of 4 with $100k Income		1.336 (1.227)			1.913* (1.130)	
Top Marginal			0.587 (0.507)			0.737 (0.470)
Observations	654,587	654,587	617,915	639,787	639,787	604,987
R^2	0.995	0.995	0.995	0.994	0.994	0.994
Adjusted R^2	0.979	0.979	0.980	0.977	0.977	0.977
Note:				*p<0.1; **p<0.05; ***p<0.01		

Estimates of (3.4). All dependent and independent variables are in terms of logs. The unit of observation is ZIP code-pair-month, meaning each ZIP code-month will appear multiple times, depending on the number of pairs of which it is a part. Coefficients should be interpreted as elasticities. All independent variables are average state net-of-tax rates on the population listed. All models include ZIP code and pair-time fixed effects. Standard errors, clustered separately at the border segment (pair of states) and state levels, are in parentheses.

code will appear multiple times—one for each border pair of which it is a part. Since each pair also is subject to serial autocorrelation, this means that observations on the same border segment may be correlated *even at different points in time*. Thus, I follow Dube, Lester and Reich (2010) and Conley and Taber (2011) and cluster separately at the border segment and state levels.

Point estimates of the elasticity of home prices with respect to the net-of-tax rate range from 0.687 to 2.327 across these specifications, all of which are large. However, the large standard errors make meaningful inference difficult. I will discuss these standard errors further in Section 3.5, but they mostly result from two phenomena: clustering at the segment and state levels throws out meaningful time series information about the dynamic response to taxes; and most of the tax changes are quite small, which is equivalent to few treated clusters. I address the former issue in Section 3.4.2 as a side effect of testing the parallel trends assumption, and the latter in Section 3.4.3.

3.4.2 Dynamic Response

The identification assumption necessary for the validity of the strategy in Section 3.4.1 is commonly known as parallel trends. That is, suppose ZIP code i is a member of one state that changes its tax rate from year t to year $t+1$, and ZIP code j is a neighboring member of another state that does not. (3.4) considers ZIP code j to be a control group for ZIP code i. The ZIP code fixed effects control for the possibility that ZIP code j is secularly (irrespective of taxes) more or less desirable than ZIP code i, but only if the gap between them would be expected[12] to have been constant in the absence of the tax change.

[12]It is important that parallel trends need only hold in expectation. Idiosyncratic violations of parallel trends won't affect the point estimate, and will just increase the standard error. To bias the result, the violations of parallel trends must correlate with the independent variable—the difference in tax rates.

The parallel trends assumption can be tested by looking for the dynamic response to the tax change. Specifically, one would expect

$$\frac{\partial \ln h_{j,t+s}}{\partial \ln(1-\tau_{jt})} \neq 0$$

iff $s \geq 0$. That is, the effect of a tax change should only be felt on or after the date of its passage.[13] Additionally, investigating the dynamic response to the tax change allows extraction of more information from the sample. Specifically, the cluster robust variance estimates (CRVEs) of the previous section merely use a single measure of the correlation of taxes and home prices within a segment, controlling for fixed effects. But the data within each cluster contains more information than that—it contains the *timing* of home price changes, and that information is useful for forming a conclusion about the home price response to taxes if it is systematically related to the timing of tax changes.

Dube, Lester and Reich (2010) suggest the following specification for estimating the dynamic effect:

$$\ln h_{ipt} = \phi_i + \psi_{pt} + \sum_{s=-(T-1)}^{T} (\eta_{-s} \Delta \ln(1-\tau)_{i,t+s}) + \eta_T \ln(1-\tau)_{i,t-T} + \mu_{ipt}. \tag{3.5}$$

As they note, specifying the tax variables in first-differences over time allows estimation of elasticities η_s of home prices in period $t+s$ with respect to a permanent tax change in period t.

Even though taxes change only yearly, I have monthly home price data, so I can estimate a monthly home price response. Choosing T presents a tradeoff; the larger T is, the better the picture of the dynamic response, but the smaller the sample: For an observation in period t to be included in the sample, the original dataset must contain

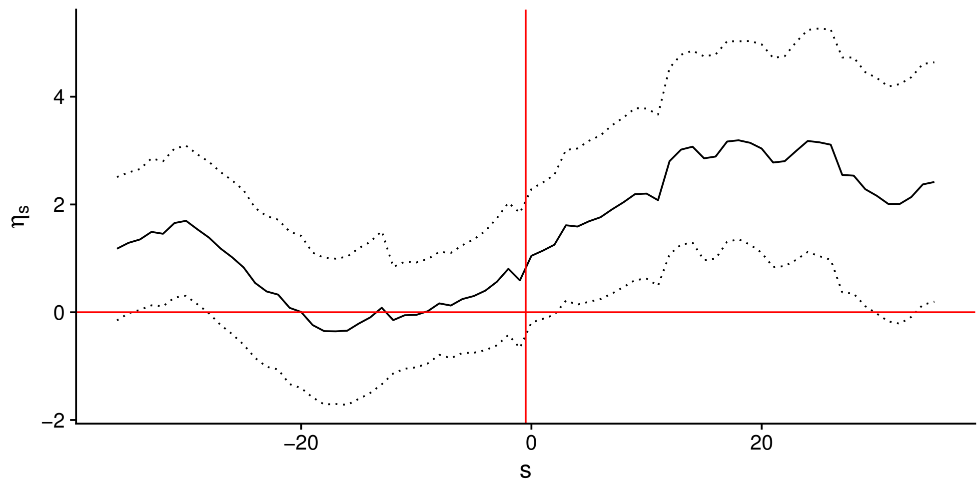
[13] or anticipated passage. I will return to this idea later.

T periods on both sides of t. I've chosen 3 years ($T = 36$) as a compromise position.

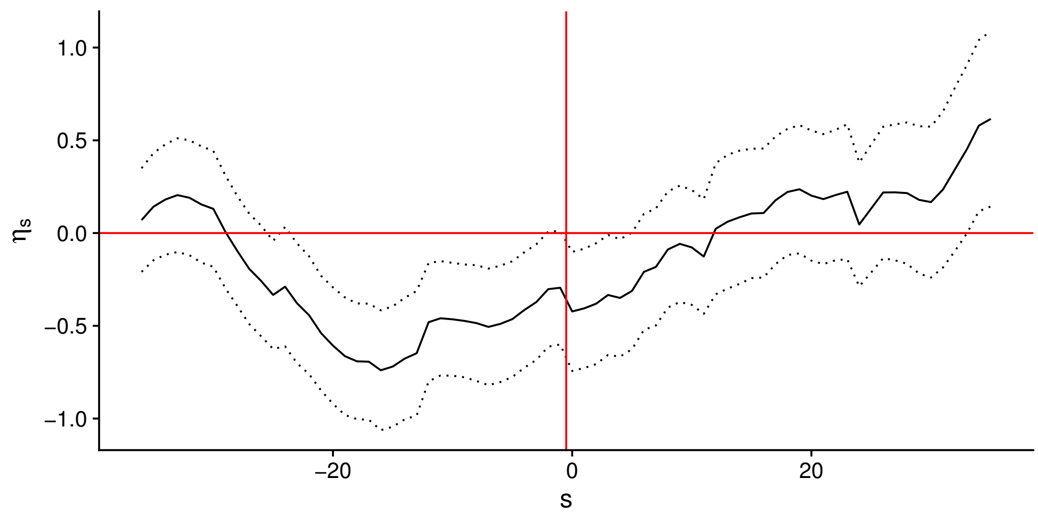
I estimate (3.5) six ways: median and per square foot prices as the dependent variable; and the three tax measures as the independent variable. I continue to use standard errors clustered separately at the state and border segment levels. Figure 3.4 visualizes these estimates for two of these specifications. Others yield qualitatively similar results and can be found in Appendix C.2. The results are inconclusive, since there is a trend upward in home prices both before and after a given tax change is implemented. Additionally, especially for the top panel, standard errors are quite wide. Nonetheless, especially in the top panel, prior to the period of the tax change, the elasticity is not statistically different from zero, whereas it becomes significantly positive shortly after the tax change; thus, a relationship may exist.

3.4.3 Event Study

One shortcoming of the data discussed in Section 3.3.3 is that most year-over-year tax changes in the sample period were quite small, with a standard deviation of only roughly 0.15%. This presents a major problem of external validity, in addition to making CRVEs unreliable (Conley and Taber, 2011). Even a sizable elasticity like 2.5, estimated in Section 3.4.2, corresponds to an increase in home values of around 0.5% for a tax cut of 0.15%. Extrapolating this to a 2.5% increase in home values for a tax cut of 1% seems unjustified. One way to examine the validity of such an extrapolation is to use an event study approach, examining those instances in which states changed their tax rates markedly relative to their neighbors. I chose the largest ten changes in the top marginal rate over this period (excluding Hawaii, and two instances in which the tax change was temporary, and withdrawn within three years) and then used maps of Zillow-covered ZIP codes along the state's borders to eliminate instances with obviously insufficient data. I focus on three representative instances (of the remaining five):



(a) Taxes on families earning $100,000, per square foot home prices



(b) Top marginal tax rates, per square foot home prices

Figure 3.4: Dynamic price responses to tax changes via estimation of (3.5). Elasticities of home prices in month $t + s$ with respect to a permanent change in the net-of-tax rate in month t are given by the solid line, with 95% confidence intervals bounded by the dotted lines. Standard errors are clustered separately at the border segment and state levels.

Figure 3.5: Data availability for ZIP codes on the New Jersey border. Shaded areas are ZIP codes with data.

- New Jersey raised top rates by 2.6 percentage points in 2004.
- North Carolina cut top rates by 2.18 percentage points in 2014.
- Illinois raised top rates by two percentage points in 2011.

New Jersey 2004. This is the famous "Millionaire's Tax" covered extensively by Young and Varner (2011). It raised the top marginal rate, affecting earnings over $500,000, by 2.6 percentage points. It was passed in June of 2004, and was effective retroactive to the beginning of the tax year. It was accompanied by an expansion of property tax credits, and therefore represents an increase in the progressiveness of the tax code more than an overall tax increase. The availability of home price data on the New Jersey border can be seen in Figure 3.5.

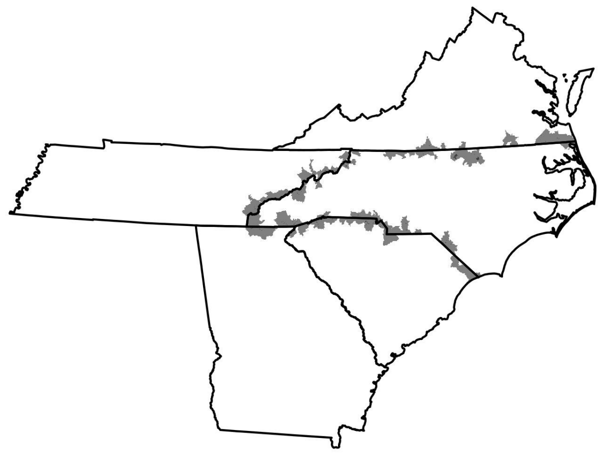


Figure 3.6: Data availability for ZIP codes on the North Carolina border. Shaded areas are ZIP codes with data.

North Carolina 2014. In August of 2013, North Carolina passed sweeping changes to its tax code (Sahadi, 2013). It reduced marginal rates on personal income substantially, with the top rate dropping by 2.18 percentage points, and lower rates dropping by smaller amounts. Additionally, standard deductions were expanded. This tax cut was passed several months ahead of when it took effect, and so might have been somewhat anticipated. The availability of home price data on the North Carolina border can be seen in Figure 3.6.

Illinois 2011. In 2011, Illinois raised income taxes by 2% across the board. The bill was passed on January 12, 2011, retroactive to January 1, 2011 (Henchman and Padgitt, 2011). According to Long (2011), the bill was passed with the state's budget under duress. It passed with an incredibly close vote and therefore was unlikely to

Figure 3.7: Data availability for ZIP codes on the Illinois border. Shaded areas are ZIP codes with data.

be fully anticipated. Illinois has no local income taxes that might confound this tax change. On the other hand, the bill provides for a phase-out, and so might be regarded as temporary by residents. Data on the Illinois borders with Indiana and Wisconsin are patchy, but there are sufficiently many data points to obtain reasonably precise estimates. The availability of home price data on the Illinois border can be seen in Figure 3.7.

To avoid issues resulting from unbalanced panels while viewing the dynamic response to the policy change, I conduct the event study by using a regression on ZIP code pairs:

$$\ln h_{ipt} = \phi_i + \psi_{pt} + \sum_{s=T_1}^{T_2} \eta_s \mathbb{1}[t = s] \times TREAT_i + \mu_{ipt}, \tag{3.6}$$

where $TREAT$ represents whether ZIP code i belongs to the treated state (New Jersey, North Carolina, or Illinois). I omit the interaction term for the month during which the tax change took place, so all ηs represent the *change* in home prices relative to the month during which the event took place.

Standard errors are challenging. Continuing to cluster errors at the segment and state level separately results in only a handful of clusters, which is not sufficient to estimate standard errors. I do cluster at the ZIP code level to correct for mechanical correlation between the error terms due to repetition of a ZIP code as a member of multiple pairs, as well as any autocorrelation of the error term.

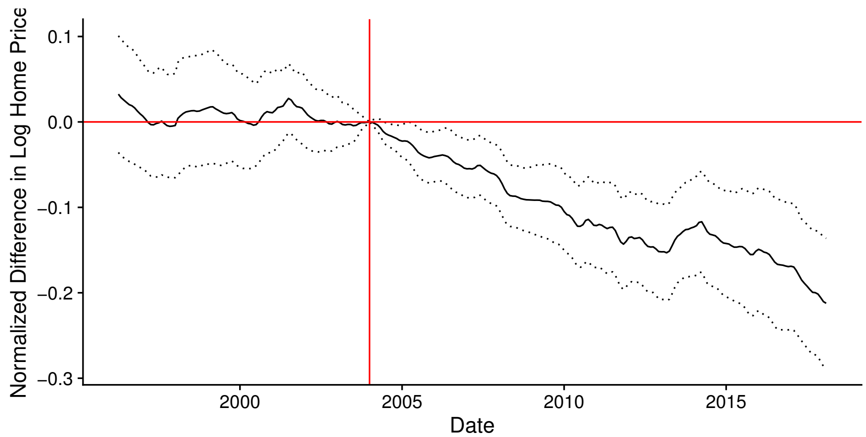
Figures 3.8, 3.9, and 3.10 depict results for New Jersey, North Carolina, and Illinois, respectively.

These results are representative of the sorts of results I obtained for other event studies. The New Jersey event study shows home prices beginning to sharply fall in concert with the rise in top tax rates; however, the magnitude is so large is to not be plausibly related to the tax change. The North Carolina event study shows no effect whatsoever. Finally, the Illinois study shows a drop in the point estimate at the time of the tax change, but the 95% confidence interval includes zero (no effect) for several months (or years) afterward, and the point estimate had already begun to drop several months prior to the tax change. As a result, the results are inconclusive on the whole.

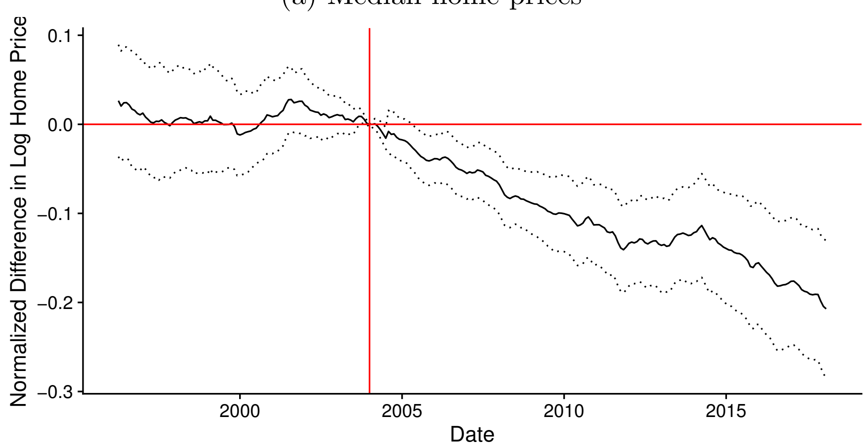
3.5 Discussion

3.5.1 Alternative Standard Errors

As mentioned, I follow Dube, Lester and Reich (2010) and use standard errors that are clustered at the border segments level and state level separately. Each dimension has over 40 clusters, making inference generally reliable. Furthermore, Bertrand, Duflo and Mullainathan (2004) show that clustering at the state level (their setting features no border pairs) is required to account for quite general patterns of serial correlation. These

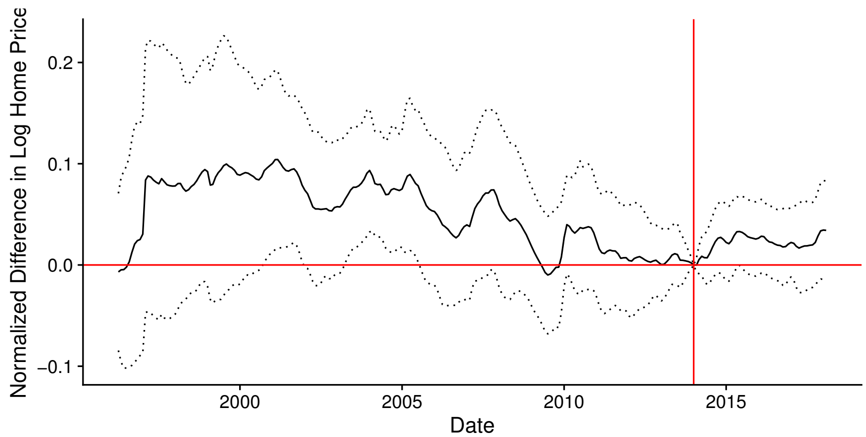


(a) Median home prices

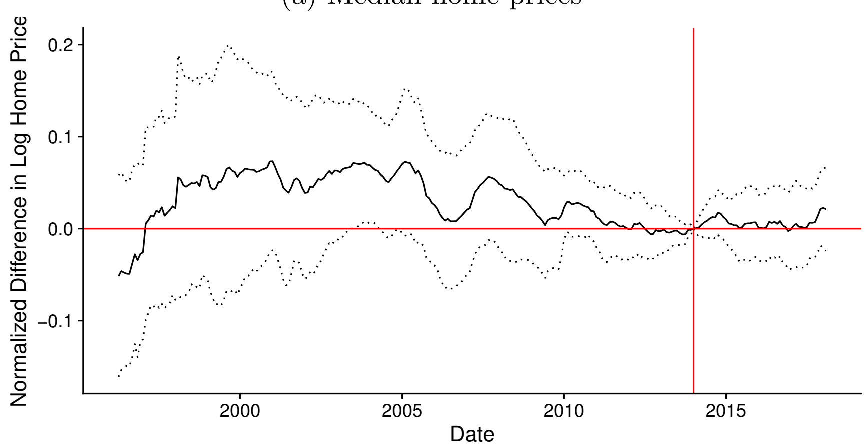


(b) Per square foot home prices

Figure 3.8: New Jersey event study. The estimate of η_s from (3.6) is given by the solid line, with 95% confidence intervals bounded by the dotted lines. These values are normalized such that the value is 0 in January 2004, the month of the tax change, marked by a vertical line. Standard errors are clustered at the ZIP code level.

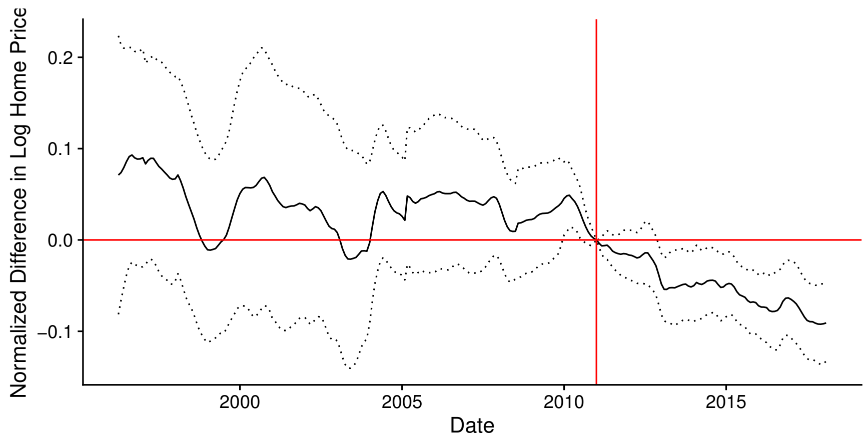


(a) Median home prices

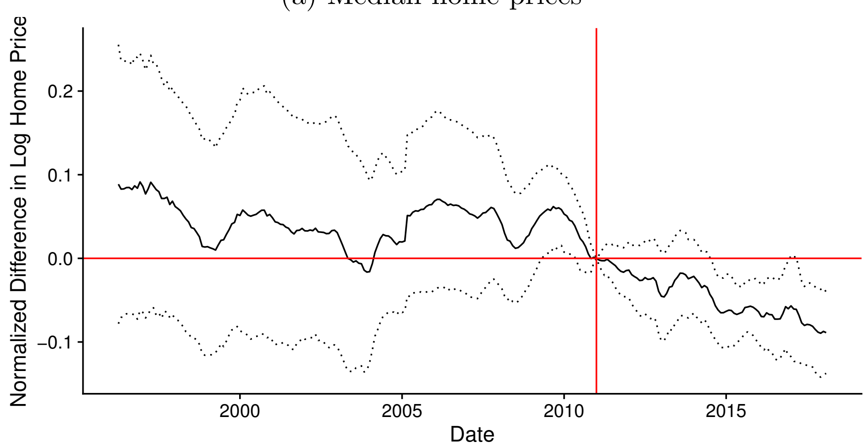


(b) Per square foot home prices

Figure 3.9: North Carolina event study. The estimate of η_s from (3.6) is given by the solid line, with 95% confidence intervals bounded by the dotted lines. These values are normalized such that the value is 0 in January 2014, the month of the tax change, marked by a vertical line. Standard errors are clustered at the ZIP code level.



(a) Median home prices



(b) Per square foot home prices

Figure 3.10: Illinois event study. The estimate of η_s from (3.6) is given by the solid line, with 95% confidence intervals bounded by the dotted lines. These values are normalized such that the value is 0 in January 2011, the month of the tax change, marked by a vertical line. Standard errors are clustered at the ZIP code level.

standard errors are, unfortunately, large enough to prohibit detection of reasonable-sized effects in many specifications.

It might appear that one could refine the sample to obtain more efficient estimates. In particular, the identification assumption required for validity of the border pair approach is that, conditional on pair-time and ZIP code fixed effects, home prices would have evolved in the same way on both sides of the border in the absence of any tax change. This is conceptually equivalent to the assumption of matched pairs, and it is standard in such studies to cluster at the pair level, which still allows for arbitrary serial correlation within the pair.

The complexity of the geographic map, however, means that the same ZIP code will appear in more than one pair. This clearly leads to mechanical correlation across all pairs that share a common ZIP code. But it is worse than that; due to the pair-time fixed effects, this can lead to a chained effect that causes mechanical correlation of *all* observations along a given border or within a given state. This leads to the two-way cluster robust variance estimate (CRVE) implemented in the previous section. This complexity can be removed, however, by ensuring that each ZIP code appears as part of only one border pair. Once this is guaranteed, there should be no mechanical correlation of the error term outside a border pair.

Table 3.5 shows the results of reestimating (3.4) using this method. In particular, I have randomly sorted the list of all adjacent pairs or ZIP codes, and then iteratively eliminated pairs until each ZIP code appears in only one pair. I cluster the standard errors at the pair level, as is typical of matched pairs tests. The point estimates are qualitatively similar, suggesting that the slimming of the sample did not meaningfully affect the estimates, but with somewhat smaller standard errors; as a result, several of the estimates are statistically significant.

However, it turns out that this method substantially over-rejects a null hypothesis.

Table 3.5: Estimates using each ZIP code exactly once, clustered at pair level

Tax Rate	Median Home Price			Per Sq. Ft. Home Price		
Family of 4 with $50k Income	1.161 (0.897)			1.928** (0.848)		
Family of 4 with $100k Income		1.043 (0.874)			1.431* (0.802)	
Top Marginal			0.517 (0.334)			0.722** (0.310)
Observations	230,193	230,193	217,317	223,585	223,585	211,441
R^2	0.994	0.994	0.994	0.994	0.994	0.994
Adjusted R^2	0.979	0.979	0.980	0.977	0.977	0.978
Note:						*p<0.1; **p<0.05; ***p<0.01

Estimates of (3.4) using a sample of border pairs such that no ZIP code appears in more than one pair. All dependent and independent variables are in terms of logs. The unit of observation is ZIP code-month. Coefficients should be interpreted as elasticities. All independent variables are average state net-of-tax rates on the population listed. All models include ZIP code and pair-time fixed effects. Standard errors, clustered at the pair level, are in parentheses.

I test this by following Bertrand, Duflo and Mullainathan (2004) and generating 1,000 placebo policies. States independently have a 50% chance of receiving a placebo treatment, and then the timing of that treatment is randomly and independently assigned. A valid statistical test of size .05 should reject the null hypothesis of zero effect of the placebo treatment about 5% of the time. This procedure, however, rejects the null over 20% of the time. Meanwhile, the two-way clustered standard errors of the previous section lead to rejection of the null for a placebo treatment about 7% of the time—much closer to the proposed size of the test.[14]

The difference in these two estimates of the standard error is allowance for the possibility of correlation of the error term across different ZIP codes on the same side of a state border, which might be attributable to statewide shocks that do not propagate beyond the state border, such as state laws. So long as these shocks are not systematically correlated with state income tax changes, they do not bias the estimate of the elasticity, but they *do* affect the standard error. Thus, the usual two-way clustered errors are most appropriate.

3.5.2 Parallel Trends

There appears to be suggestive evidence that home prices may be hurt by increased state income taxes. However, dynamic results for many measures of taxes call into question the parallel trends assumption necessary for correctly interpreting a difference-in-differences estimate as causal. Many specifications presented in Section 3.4.2 and Appendix C.2 find that home prices are positively associated with tax cuts even two years before the tax cut in question. Thus, one might question whether the border pair partner is a suitable control group, since the two home prices begin to separate even

[14]This analysis was performed on an earlier version of the dataset; it is quite computationally intensive, and so I did not repeat it on the newer version of the dataset.

prior to treating half of the pair with a new tax rate.

The violation of parallel trends should not be interpreted as evidence *against* the hypothesis that increased state income taxes hurt home prices. Many stories may explain such a pattern. One is that tax changes are anticipated and capitalized in home prices substantially before they take effect, in which case the causal interpretation would still be valid. On the other hand, another possibility is that secular economic conditions are specific to states, even near the border, and drive both home prices and taxes—in opposite directions; this story would invalidate any causal interpretation.

3.5.3 Other Mechanisms and Confounding Factors

These estimates can only be seen as causal to the extent that, conditional on ZIP code fixed effects and pair-time fixed effects, the error term—which captures changes in relative desirability for reasons other than taxes—is uncorrelated with changes in taxes. While the border pair partner likely controls well for general economic conditions in the area, it does *not* control for other policies (besides taxes) that depend on one's state of residence, such as public goods provision. If government budgets are balanced in the long run, then innovations to the path of future labor taxes are likely to be correlated with innovations to the path of public goods, which would violate this assumption. However, such a state of affairs would bias my results toward zero; as taxes rise, public goods would also rise, making it more desirable to reside in that state, and driving up home prices there—the exact opposite of the effect I find.

I attempted to address a different concern—that the results are driven by implausible responses to tiny tax changes—using 3 event studies: a New Jersey tax increase in 2004, a North Carolina tax cut in 2014, and an Illinois tax increase in 2011. The results from the event study exercise were mixed. North Carolina results provided little evidence of the response of home prices to taxes. On the other hand, New Jersey results are a textboook example of event study evidence of causal effects: There are no strong pre-trends, a sharp change at the time of the tax increase, and statistical significance after the event. Illinois results lay somewhere in between.

A few other concerns deserve mention. One is property taxes, another ingredient in the overall cost of living in one state versus another. Feldstein and Wrobel (1998) explicitly accounts for property taxes, but only by assuming that property values in a state are constant over the sample period, and backing out the implied millage rates from property tax collections. Since my study focuses precisely on changing property values, this method is clearly not at my disposal. However, this omission only matters to the extent that changes in property tax millage rates correlate with changes in income tax millage rates—an empirical question I do not take up here but is worthy of further investigation.

Second, these estimates, to the extent they are accepted as causal, should for some reasons be seen as upper bounds on the elasticity of home prices with respect to the state net-of-income-tax rate, and for other reasons as lower bounds. At the border, homeowners have an obvious choice when taxes in their home state rise—they can sell their homes and buy others just across the border. In a ZIP code in the center of a state, however, escaping the higher state income taxes is not so simple of a proposition; homeowners would be required to relocate to another metropolitan area and probably search for a new job. For this reason, my border estimates should be seen as upper bounds on the true elasticity across the country.

On the other hand, if a worker works in one state and resides in another, he typically pays the higher of the two state income taxes. Thus, workers employed in the higher tax state but living in the lower tax state will face no incentive to relocate if their state of residence adjusts its income tax rate up or down. For this reason, my estimates should be seen as lower bounds on the true elasticity across the country.[15]

Finally, Coglianese (2015) questions the validity of border pair research designs in general. He shows that, with respect to employment rates, trends in the rest of the state are predictive of a county's employment situation, even after controlling for the situation in a bordering county in a neighboring state, and even in states with no change in the length of unemployment insurance benefits.

3.5.4 Magnitudes and MVPF Calculations

Temporarily putting aside the concerns discussed in the previous subsection, I consider whether the estimated elasticities have a plausible magnitude. To do so, I perform a back-of-the envelope calculation. For a family earning \$50,000 post-tax, a one percent drop in net-of-tax wages results in a loss of income of \$500 per year. Meanwhile, a home valued at \$500,000 and suffering a one percent drop in value results in a reduction in value of \$5,000. Assuming an interest-only loan at 5% yields a drop in annual interest payments of \$250. Thus, an elasticity around 2, the top end of the estimates found in the empirical section, would appear to be roughly the right magnitude, as it would leave the disposable (after housing) income of the family unchanged.

Finally, I return to (3.2), which I reproduce here,

$$MVPF = \frac{\frac{1}{\tau^L}R^L - \frac{\eta}{1-\tau^L}R^H}{\frac{1}{\tau^L}R^L - \frac{1}{1-\tau^L}\left[\eta R^H + \left(\underbrace{e_L^w + \eta e_L^H}_{\text{Total response}}\right)R^L + \left(\underbrace{e_H^w + \eta e_H^h}_{\text{Total response}}\right)R^H\right]}$$

with the goal of demonstrating the magnitude by which the MVPF may be mismeasured if the home price effect is not considered. Recall that this can be compared to the

[15]Migration is not the only mechanism by which taxes might affect home prices. For example, if everyone in the country saw their taxes increase substantially, demand would be expected to drop purely due to income effects, which would cause home prices to fall.

simple MVPF formula

$$MVPF = \frac{1}{1 - \frac{\tau^L}{1-\tau^L}(e_L^w + \eta e_L^h)}$$

where $e_L^w + \eta e_L^h$ is the total labor supply response to the tax change, and the house price response is ignored.

I calibrate this as follows. The state labor tax rate is around 6% (Kaeding, 2016). The ratio of state labor tax revenues to property tax revenues ranges from 4/7 to 11/7 (Malm and Kant, 2013).[16] Studies estimate the total response to federal income tax changes as having elasticities between 0.33 and 2 (Chetty et al., 2011); I assume the total response to state taxes is similar—which it would be if the labor supply response to home prices is small, a believable assumption.

With these assumptions, the MVPF using the simple formula that does not account for GE effects on home prices ranges from 1.02[17] to 1.15.[18] The properly calculated MVPF further depends on η, the elasticity estimated in this paper, and $e_H^w + \eta e_H^h$, the total response of housing consumption to the change in the labor tax rate—both the direct response to the drop in net-of-tax wages, and the indirect response to the drop in home prices—a value very difficult to estimate. I assume that the housing consumption response is nonpositive,[19] and η lies between 0.5 and 2.5, based on the results in the

[16]First, it is unclear whether local revenue loss should be included—the agency making most decisions about labor income taxes (the state government) is different from that obtaining most revenue from property taxes (local government). However, it seems reasonable to consider local governments to be agents of the state government. Second, labor income and property taxes together account for only about 55% of state and local revenue, with most of the rest coming from sales taxes. The 4/7 figure ignores sales taxes, since in border regions they are often paid by nonresidents; the 11/7 figure counts sales taxes as labor income taxes in the public finance tradition.

[17]Labor elasticity of 0.33

[18]Labor elasticity of 2

[19]That is, I assume that the net effect of an increase in labor taxes and a decrease in home prices *does not* lead to *higher* housing consumption.

foregoing sections. Then a lower bound on the MVPF ranges from 1.02[20] to 1.22[21]. This 1.22 estimate is quite different from the upper bound of 1.15 calculated using the simple MVPF formula, and suggests that ignoring the housing price response when formulating policy is quite dangerous.[22]

3.6 Conclusion

In summary, I find suggestive but inconclusive evidence that home prices are responsive to state income taxes, with suggested estimates of the elasticity of home prices with respect to net-of-tax rates of between 0.5 and 2.5. The fact that some specifications are less conclusive highlights the importance of checking border pair designs for robustness using dynamic specifications and event studies. Calibrating a formula for the marginal value of public funds for state income tax adjustments shows that ignoring the effect on home prices can lead to very erroneous results, meaning that obtaining a good estimate of this elasticity is important, and worthy of further study.

I see three major directions for future work. First, one could analyze the response of home prices to other relevant taxes—for example, property taxes or local income taxes (in those jurisdictions that have them). Second, one could attempt to provide more conclusive evidence of the effect suggested by this paper by using transaction-level data. This would not directly change standard errors; the clustering at the border segment level means that more large tax change events would be required to increase precision, not merely more observations per event. However, transaction level data would allow

[20]Labor elasticity of 0.33, $R^L/R^H = 11/7$, $\eta = 0.5$

[21]Labor elasticity of 2, $R^L/R^H = 4/7$, $\eta = 2.5$

[22]This could be substantially larger if housing consumption drops as a result of the tax. For example, if $e_H^w + \eta e_H^h = 1$—that is, a 1% reduction in the net-of-tax rate leads to a 1% reduction in amount of housing consumed—then this range shifts up substantially.

finer observation of location, perhaps leading to a regression discontinuity approach at the state border, rather than the broader ZIP code pair approach. This would allow for a better quasi-control group, and make it more likely that conclusive causal evidence would be uncovered.

Third, in the MVPF calibration I undertook in Section 3.5, I was forced to make two assumptions regarding empirically estimable quantities not investigated here. First, I assumed that the labor supply response to state income taxes and federal income taxes is similar, which is equivalent to assuming that the labor supply response to lower home prices is small. Second, I assumed that the total response of housing consumption to state income taxes—both the direct response, and the countervailing response to the ensuing lower home prices—is weakly negative, and considered an overall elasticity of 0 and 1. Both of these assumptions could be empirically investigated, and the MVPF calculation could be updated with more precise values.

References

Aiyagari, S Rao. 1995. "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting." *Journal of Political Economy*, 103(6): 1158–1175.

Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppala. 2002. "Optimal Taxation without State-Contingent Debt." *Journal of Political Economy*, 110(6): 1220–1254.

Altig, David, Alan J Auerbach, Laurence J Kotlikoff, Kent A Smetters, and Jan Walliser. 2001. "Simulating Fundamental Tax Reform in the United States." *The American Economic Review*, 91(3): 574–595.

Anacker, Katrin B. 2010. "Still Paying the Race Tax? Analyzing Property Values in Homogeneous and Mixed-Race Suburbs." *Journal of Urban Affairs*, 32(1): 55–77.

Atkinson, A B, and A Sandmo. 1980. "Welfare Implications of the Taxation of Savings." *The Economic Journal*, 90(359): 529–549.

Auerbach, Alan J., and Laurence J. Kotlikoff. 1987. *Dynamic Fiscal Policy.* Cambridge:Cambridge University Press.

Autor, David H, and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *The American Economic Review*, 103(5): 1553–1597.

Badger, Emily. 2016. "'This Can't Happen by Accident'." *The Washington Post*, 2016(May 2).

Bailey, Martin J. 1966. "Effects of Race and of Other Demographic Factors on the Values of Single-Family Homes." *Land Economics*, 42(2): 215.

Ball, Laurence, and N. Gregory Mankiw. 2007. "Intergenerational Risk Sharing in the Spirit of Arrow, Debreu, and Rawls, with Applications to Social Security Design." *Journal of Political Economy*, 115(4): 523–547.

Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy*, 87(5, Part 1): 940–971.

Bayer, Patrick, Marcus Casey, Fernando Ferreira, and Robert Mcmillan. 2013. "Estimating Racial Price Differentials in the Housing Market Estimating Racial Price Differentials in the Housing Market." *Economic Research Initiatives at Duke Working Paper Series*, 142.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *The American Economic Review*, 94(4): 991–1013.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Blanchard, Olivier J. 1985. "Debt, Deficits, and Finite Horizons." *The Journal of Political Economy*, 93(2): 223–247.

Bound, John, and Richard B Freeman. 1992. "What Went Wrong? The Erosion of Relative Earnings and Employment Among Young Black Men in the 1980s." *The Quarterly Journal of Economics*, 107(1): 201–232.

Cain, Glen G. 1986. "The economic analysis of labor market discrimination: A survey." In *Handbook of Labor Economics*. Vol. 1, , ed. David Card and Orley Ashenfelter, Chapter 13, 693–785. North Holland.

Canner, Glenn B., Dolores S. Smith, Nancy E. Bowen, and Florence M. Benkoivic. 1991. "Home Mortgage Disclosure Act: Expanded Data on Residential Lending." *Federal Reserve Bulletin*, 77.

Center on Budget and Policy Priorities. 2017. "Policy Basics: Where Do Our Federal Tax Dollars Go?"

Cervero, Robert, and Michael Duncan. 2004. "Neighbourhood Composition and Residential Land Prices: Does Exclusion Raise or Lower Values?" *Urban Studies*, 41(2): 299–315.

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica*, 54(3): 607–622.

Chari, V.V., and Patrick J. Kehoe. 1999. "Optimal fiscal and monetary policy." *Handbook of Macroeconomics*, 1: 1671–1745.

Chari, V V, Lawrence J Christiano, and Patrick J Kehoe. 1994. "Optimal Fiscal Policy in a Business Cycle Model." *Journal of Political Economy*, 102(4): 617–652.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics*, 1(1): 451–488.

Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2018. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *NBER Working Paper Series*, 24441.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *The Quarterly Journal of Economics*, 126(4): 1593–1660.

Choudhury, Sharmila. 2002. "Racial and Ethnic Differences in Wealth and Asset Choices." *Social Security Bulletin*, 64(2).

Coate, Douglas, and James Vanderhoff. 1993. "Race of the Homeowner and Appreciation of Single-family Homes in the United States." *Journal of Real Estate & Economics*, 7(3): 205–212.

Coate, Douglas, and Richard W Schwester. 2011. "Black-White Appreciation of Owner-Occupied Homes in Upper Income Suburban Integrated Communities: The Cases of Maplewood and Montclair, New Jersey." *Journal of Housing Research*, 20(2): 127–139.

Coglianese, John. 2015. "Do Unemployment Insurance Extensions Reduce Employment?"

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing Capital? Not a Bad Idea After All!" *American Economic Review*, 99(1): 25–48.

Conley, Timothy G., and Christopher R. Taber. 2011. "Inference with Difference in Differences with a Small Number of Policy Changes." *Review of Economics and Statistics*, 93(1): 113–125.

Cutler, David M. 1988. "Tax Reform and the Stock Market: An Asset Price Approach." *American Economic Review*, 78(5): 1107–17.

Dee, Thomas S. 2005. "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *The American Economic Review*, 95(2, Papers and Proceedings): 158–165.

Diamond, Peter A. 1965. "National Debt in a Neoclassical Growth Model." *American Economic Review*, 55(4): 1126–1150.

Diamond, Peter A, and James A Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Eficiency." *The American Economic Review*, 61(1): 8–27.

Diamond, Rebecca, and Tim McQuade. 2018. "Who Wants Affordable Housing in their Backyard? An Equilibrium Analysis of Low Income Property Development." *Journal of Political Economy*, Forthcoming.

Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties." *Review of Economics and Statistics*, 92(4): 945–964.

Eissa, Nada. 1995. "Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment." *NBER Working Paper Series*, 5023.

Eligon, John, and Robert Gebeloff. 2016. "Affluent and Black, and Still Trapped by Segregation."

Erosa, Andres, and Martin Gervais. 2002. "Optimal Taxation in Life-Cycle Economies." *Journal of Economic Theory*, 105(2): 338–369.

Erosa, Andres, and Martin Gervais. 2012. "On the optimality of age-dependent taxes and the progressive U.S. tax system." *Journal of Economic Dynamics and Control*, 36(4): 682–691.

Escolano, Julio. 1992. "Optimal Taxation in Overlapping Generations Models." *Mimeo.*

Farhi, Emmanuel. 2010. "Capital Taxation and Ownership When Markets Are Incomplete." *Journal of Political Economy*, 118(5): 908–948.

Farhi, Emmanuel, and Iván Werning. 2010. "Progressive Estate Taxation." *Quarterly Journal of Economics*, 125(2): 635–673.

Farhi, Emmanuel, and Iván Werning. 2013. "Insurance and Taxation over the Life Cycle." *The Review of Economic Studies*, 80(2): 596–635.

Farhi, Emmanuel, Christopher Sleet, Iván Werning, and Sevin Yeltekin. 2012. "Non-linear Capital Taxation Without Commitment." *The Review of Economic Studies*, 79(4): 1469–1493.

Feenberg, Daniel, and Elisabeth Coutts. 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management*, 12(1): 189.

Feldstein, Martin. 1995. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act." *Journal of Political Economy*, 103(3): 551–572.

Feldstein, Martin, and Marian Vaillant Wrobel. 1998. "Can state taxes redistribute income?" *Journal of Public Economics*, 68(3): 369–396.

Feldstein, Martin S. 1998. *Privatizing social security.* University of Chicago Press.

Flippen, Chenoa. 2004. "Unequal Returns to Housing Investments? A Study of Real Housing Appreciation among Black, White, and Hispanic Households." *Social Forces*, 82(4): 1523–1551.

Fortin, Bernard, Thomas Lemieux, and Pierre Frechette. 1994. "The Effect of Taxes on Labor Supply in the Underground Economy,'." *American Economic Review*, 84(1): 231–254.

Friedman, Jerome H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *The Annals of Statistics*, 29(5): 1189–1232.

Friedman, John N. 2009. "The incidence of the Medicare prescription drug benefit: using asset prices to assess its impact on drug makers." *Working paper.*

Garriga, Carlos. 2017. "Optimal Fiscal Policy in Overlapping Generations Models." *Working Paper*, 2017-032A.

Goolsbee, Austan. 2000. "What Happens When You Tax the Rich? Evidence from Executive Compensation." *Journal of Political Economy*, 108(2): 352–378.

Gordon, Roger H., and Hal R. Varian. 1988. "Intergenerational risk sharing." *Journal of Public Economics*, 37(2): 185–202.

Green, Jerry. 1977. "Mitigating Demographic Risk Through Social Insurance." *NBER Working Paper Series*, 215.

Grodsky, Eric, and Devah Pager. 2001. "The Structure of Disadvantage: Individual and Occupational Determinants of the Black-White Wage Gap." *American Sociological Review*, 66(4): 542.

Guren, Adam M. 2018. "House Price Momentum and Strategic Complementarity." *Journal of Political Economy*, Forthcomin(Forthcoming).

Hardt, Moritz, Eric Price, and Nati Srebro. 2016. "Equality of Opportunity in Supervised Learning."

Harenberg, Daniel, and Alexander Ludwig. 2014. "Social Security and the Interactions Between Aggregate and Idiosyncratic Risk Social Security and the Interactions Between Aggregate and Idiosyncratic Risk." *ETH Risk Center—Working Paper Series.*

Harris, David R. 1999. ""Property Values Drop When Blacks Move in, Because...": Racial and Socioeconomic Determinants of Neighborhood Desirability." *American Sociological Review*, 64(3): 461.

Harrison, Roderick J., and Claudette E. Bennett. 1995. "State of the Union: America in the 1990s." In *State of the Union: America in the 1990s.* , ed. Reynolds Farley, Chapter 4, 141—-210. New York, NY:Russell Sage Foundation.

Henchman, Joseph, and Kail Padgitt. 2011. "Illinois Approves Sharp Income Tax Increase, Fourth-Highest Corporate Tax Rate." *Tax Foundation*, 256.

Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy*, 30(1): 51–89.

Jan, Tracy. 2017. "White families have nearly 10 times the net worth of black families. And the gap is growing. - The Washington Post." *The Washington Post*, 2017(September 28).

Jones, Janelle. 2017. "The racial wealth gap: How African-Americans have been shortchanged out of the materials to build wealth."

Judd, Kenneth L. 1985. "Redistributive taxation in a simple perfect foresight model." *Journal of Public Economics*, 28(1): 59–83.

Judd, Kenneth L. 1998. *Numerical Methods in Economics.* Cambridge, MA:The MIT Press.

Judd, Kenneth L, Lilia Maliar, Serguei Maliar, and Rafael Valero. 2014. "Smolyak method for solving dynamic economic models: Lagrange interpolation, anisotropic grid and adaptive domain." *Journal of Economic Dynamics and Control*, 44: 92–123.

Kaeding, Nicole. 2016. "State Individual Income Tax Rates and Brackets for 2016." *The Tax Foundation.*

Kaplow, Louis. 2003. "Concavity of Utility, Concavity of Welfare, and Redistribution of Income." *John M. Olin Center for Law, Economics, and Business Discussion Paper*, 437.

Kilbourne, Barbara, Paula England, and Kurt Beron. 1994. "Effects of Individual, Occupational, and Industrial Characteristics on Earnings: Intersections of Race and Gender." *Social Forces*, 72(4): 1149.

Kim, Sunwoong. 2000. "Race and Home Price Appreciation in Urban Neighborhoods: Evidence from Milwaukee, Wisconsin." *The Review of Black Political Economy*, Fall 2000.

Kotlikoff, Laurence J, Kent Smetters, and Jan Walliser. 1999. "Privatizing Social Security in the United States–Comparing the Options*." *Review of Economic Dynamics*, 2: 532–574.

Krueger, Dirk, and Felix Kubler. 2004. “Computing equilibrium in OLG models with stochastic production.” *Journal of Economic Dynamics and Control*, 28(7): 1411–1436.

Krueger, Dirk, and Felix Kubler. 2006. “Pareto-Improving Social Security Reform when Financial Markets Are Incomplete!?” *American Economic Review*, 96(3): 737–755.

Krusell, Per, and Anthony A. Smith, Jr. 1998. “Income and Wealth Heterogeneity in the Macroeconomy.” *Journal of Political Economy*, 106(5): 867–896.

Lake, Robert W. 1979. “Racial Transition and Black Homeownership in American Suburbs.” *The ANNALS of the American Academy of Political and Social Science*, 441(1): 142–156.

Linden, Leigh L., and Jonah E. Rockoff. 2008. “There Goes the Neighborhood? Estimates of the Impact of Crime Risk on Property Values From Megan’s Laws.” *American Economic Review*, 98(3): 1103–1127.

Long, James E, and Steven B Caudill. 1992. “Racial Differences in Homeownership and Housing Wealth, 1970–1986.” *Economic Inquiry*, 30(1).

Long, Ray. 2011. “Clout Street: Quinn on tax hike: ‘Our fiscal house was burning’.” *Chicago Tribune*.

Lucas, Robert E., and Nancy L. Stokey. 1983. “Optimal fiscal and monetary policy in an economy without capital.” *Journal of Monetary Economics*, 12(1): 55–93.

Macpherson, David A., and G. Stacy Sirmans. 2001. “Neighborhood Diversity and House-Price Appreciation.” *The Journal of Real Estate Finance and Economics*, 22(1): 81–97.

Malm, Liz, and Ellen Kant. 2013. “The Sources of State and Local Tax Revenues.” *The Tax Foundation*.

Manson, Steven, Jonathon Schroeder, David Van Riper, and Steven Ruggles. 2017. “IPUMS National Historical Geographic Information System: Version 12.0 [Database].” *Minneapolis: University of Minnesota.* *http://doi.org/10.18128/D050.V12.0*.

Mark, Nelson C. 2004. “Notes on Numerical Techniques for Solving Dynamic Economic Models.”

Massey, Douglas S, and Nancy A Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass.* Cambridge, MA:Harvard University Press.

Mauro, Paolo, Rafael Romeu, Ariel Binder, and Asad Zaman. 2015. "A modern history of fiscal prudence and profligacy." *Journal of Monetary Economics*, 76: 55–70.

McKernan, Signe-Mary, Caroline Ratcliffe, Eugene Steuerle, and Sisi Zhang. 2013. "How Have Wealth Inequality and Income Inequality Changed Over Time?" Urban Institute, Washington, D.C.

Miller, Claire Cain. 2015. "When Algorithms Discriminate."

Miranda, Mario J., and Paul R. Fackler. 1997. "Lecture Notes in Computational Economic Dynamics."

Mirrlees, J A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *The Review of Economic Studies*, 38(2): 175–208.

Oliver, Melvin, and Thomas Shapiro. 2006. *Black Wealth / White Wealth: A New Perspective on Racial Inequality.* . 2nd Editio ed., New York, NY:Routledge.

Piketty, Thomas, and Emmanuel Saez. 2013. "A Theory of Optimal Inheritance Taxation." *Econometrica*, 81(5): 1851–1886.

Quercia, Roberto G., George W. McCarthy, Rhonda M. Ryznar, and Aye Can Talen. 2000. "Spatio-Temporal Measurement of House Price Appreciation in Underserved Areas."

Ramirez, Edith, Julie Brill, Maureen K. Ohlhausen, Joshua D. Wright, and Terrell McSweeny. 2014. "In the Matter of CoreLogic, Inc., a corporation."

Roland G., Fryer. 2011. "Racial inequality in the 21st century: the declining significance of discrimination." *Handbook of Labor Economics*, 4: 855–971.

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2(3): 180–212.

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature*, 50(1): 3–50.

Sahadi, Jeanne. 2013. "North Carolina's Republican tax experiment." *CNN Money*, August 8,.

Samuelson, Paul A. 1958. "An exact consumption-loan model of interest with or without the social contrivance of money." *The journal of political economy*, 66(6): 467–482.

Sariyar, Murat, and Andreas Borg. 2010. "The RecordLinkage Package: Detecting Errors in Data." *The R Journal*, 2(2).

Semega, Jessica L, Kayla R Fontenot, and Melissa A Kollar. 2017. "Income and Poverty in the United States: 2016." U.S. Census Bureau: Economics and Statistics Administration.

Sherman, Michael, and Saskia le Cessie. 1997. "A comparison between bootstrap methods and generalized estimating equations for correlated outcomes in generalized linear models." *Communications in Statistics - Simulation and Computation*, 26(3): 901–925.

Shin, Laura. 2015. "The Racial Wealth Gap: Why A Typical White Household Has 16 Times The Wealth Of A Black One." *Forbes*, 2015(March): 26.

Sinai, Todd, Nicholas S Souleles, and Todd Slnai. 2005. "Owner-Occupied Housing as a Hedge against Rent Risk." *The Quarterly Journal of Economics*, 120(2): 763–789.

Smith, James P. 1995. "Racial and Ethnic Differences in Wealth in the Health and Retirement Study." *The Journal of Human Resources*, 30: S158.

Smolyak, S. 1963. "Quadrature and interpolation formulas for tensor products of certain classes of functions." *Soviet Math. Dokl.*, 4: 240–243.

Stantcheva, Stefanie. 2016. "Optimal Taxation and Human Capital Policies over the Life Cycle." *Journal of Political Economy*, Forthcomin.

The Harvard Joint Center for Housing Studies. 2015. "Improving America's Housing 2015—Emerging Trends in the Remodeling Market." Harvard University.

U.S. Census Bureau. 2017. "HINC-06. Income Distribution to $250,000 or More for Households." *https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-hinc/hinc-06.2016.html*.

Weinzierl, Matthew. 2011. "The Surprising Power of Age-Dependent Taxes." *The Review of Economic Studies*, 78(4): 1490–1518.

Werning, Iván. 2007. "Optimal Fiscal Policy with Redistribution." *The Quarterly Journal of Economics*, 122(3): 925–967.

Williams, Roberton. 2012. "What is the breakdown of revenue among federal, state, and local governments?" *Tax Policy Center: The Tax Policy Briefing Book*.

Wilson, William J. 2012. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* . Second Edi ed., Chicago:University of Chicago Press.

Wolff, Edward N. 1998. "Recent Trends in the Size Distribution of Household Wealth Data Sources and Methods." *Journal of Economic PerspectivesVolume*, 12(3): 131–150.

Young, Cristobal, and Charles Varner. 2011. "Millionaire Migration and the State Taxation of Top Incomes: Evidence from a Natural Experiment." *National Tax Journal*, 64(2): 255–83.

Appendix A

Appendix to Chapter 1

A.1 Proofs from Section 1.2

A.1.1 Formal Definition of Equilibrium

Equilibrium definitions are very similar for the two models, so I explicitly define it only for complete markets.

Definition A.1.1 (Policy Equilibrium—Complete Markets) *A policy equilibrium in the complete markets model is a collection of*

- *policies* $T_0^o, \{T(s^t), \tau^L(s^t), \tau^K(s^{t+1})\}_{t\geq 0}$
- *prices* $\{q(s^{t+1}), w(s^t), R^K(s^t)\}_{t\geq 0}$
- *an allocation* $\{c^o(s^t), c^y(s^t), \ell(s^t), k(s^t), z(s^t)\}_{t\geq 0}$
- *and government debt* $\{b(s^{t+1})\}_{t\geq 0}$

such that at all histories s^t*,* $t \geq 0$

- *The resource constraint is satisfied:*

$$c^y(s^t) + c^o(s^t) + g(s_t, t) + k(s^t) \leq F(k(s^{t-1}), \ell(s^t); A(s_t, t)) \tag{A.1}$$

- *The government's budget constraint (1.1) is satisfied*

- *The household's budget constraint (1.5) and bequest constraint (1.3)are satisfied*

- *Households optimize subject to their budget constraint, taking prices and policies as given*

- *Firms optimize, taking prices as given*

- *The state-contingent asset markets clear:*

$$b(s^t) + R^K(s^t)k(s^{t-1})(1 - \tau^K(s^t)) = c^o(s^t) + z(s^t)$$

- *The markets for capital and labor clear*

- *The no-arbitrage condition between capital and state-contingent assets holds:*

$$\sum_{s^{t+1} \succeq s^t} q(s^{t+1})R^K(s^{t+1})(1 - \tau^K(s^{t+1})) = 1$$

- *The initial old consume or bequeath their untaxed assets plus any transfer:*

$$c^o(s_0) + z(s_0) = (1 - \tau^K(s_0))R^K(s_0)k_{-1} + b(s_0) + T^o(s_0)$$

The only changes for the incomplete markets model are that households and the government face different budget constraints, and we lose the no-arbitrage condition (there are no redundant assets) and the asset market clearing condition (which is found in the household's old-age budget constraint).

A.1.2 Proof of Proposition 1.2.1

I begin by proving the "if" direction; that is, any allocation satisfying the conditions is implementable. I prove this by construction.

First, define prices

$$
\begin{aligned}
q(s^{t+1}) &= \beta Pr(s_{t+1}|s^t)\frac{u_c^o(s^{t+1})}{u_c^y(s^t)} \\
w(s^t) &= F_L(s^t) \\
R^K(s^t) &= F_K(s^t)
\end{aligned}
$$

and government debt

$$
b(s^t) = c(s^t) + z(s^t) - (1-\tau^K(s^t))R^K(s^t)k(s^{t-1}).
$$

Then define policies

$$
\begin{aligned}
\tau^L(s^t) &= 1 + \frac{u_\ell^y(s^t)}{w(s^t)u_c^y(s^t)}\ell(s^t) + \beta\mathbb{E}_t\left[\frac{u_c^o(s^{t+1})}{u_c^y(s^t)}(c^o(s^{t+1}) + z(s^{t+1})\right] \\
T_0^o &= c^o(s_0) + z(s_0) - b(s_0) \\
\tau^K(s^{t+1}) &= 1 - \frac{u_c^y(s^t)}{\beta\mathbb{E}_t[R^K(s^{t+1})u_c^o(s^{t+1})]} \\
\tau^K(s_0) &= 1
\end{aligned}
$$

Several conditions of equilibrium have already been satisfied: The resource constraint is satisfied by assumption, firms are satisfying their first order conditions at market-clearing levels of capital and labor, and the state-contingent asset markets clear by construction. Additionally, the budget constraint of the initial old is satisfied by construction. Next, the no-arbitrage condition is clearly satisfied:

$$
\sum_{s_{t+1}} \beta Pr(s_{t+1}|s^t)\frac{u_c^o(s^{t+1})}{u_c^y(s^t)}R^K(s^{t+1})\left[\frac{u_c^y(s^t)}{\beta\mathbb{E}_t[R^K(s^{t+1})u_c^o(s^{t+1})]}\right] = 1
$$

It remains to be verified that households and the government are satisfying their budget constraints, and that households are optimizing. The household budget constraint is easily verified by dividing the implementability condition by $u_c^y(s^t)$ and substituting in the prices and policies as defined. Subtracting the household budget constraint (which is satisfied with equality) from the resource constraint then yields the government budget constraint. The asset prices imply that the household's intertemporal first order conditions (Euler equations) are satisfied, while the net-of-tax wage implies that the household's intratemporal first order condition (labor-leisure tradeoff) is satisfied. Finally, the consumption-bequest tradeoff is satisfied since either the first order condition between the two is satisfied at equality,[1] or consumption is more valued than bequests, but the bequest is already at its minimum—a complementary slackness condition.

Next I continue to the "only if" direction; that is, any implementable allocation must satisfy these conditions. This is easier to show. By definition, implementable allocations must satisfy the resource constraint and the constraint on the initial old. Prices $q(s^t)$ and net-of-tax wages $(1-\tau^L(s^t))w(s^t)$ must be defined as above for households to satisfy their first order conditions. Substituting those into the household budget constraint and dividing by $u_c^y(s^t)$ yields the implementability condition. Likewise, the optimal bequest conditions come directly from the first order condition with respect to $z(s^t)$ in the household's problem.

A.1.3 Proof of Proposition 1.2.2

The proof proceeds similarly to the complete markets case, substituting prices out of budget constraints using first order conditions, and is thus omitted.

[1]To see this, take the envelope condition from (1.4)

A.2 Proofs From Section 1.3

A.2.1 Proof of Proposition 1.3.1

All of these equations are obtained through straightforward rearrangement of the planner's first order conditions.

A.2.2 Proof of Proposition 1.3.2

The implementability condition on the generation born at date t takes the form

$$1 + \beta - v'(\ell_t)\ell_t \geq 0.$$

So long as the optimum does not feature lump-sum transfers, this implementability condition holds with equality, which means that labor is fixed at ℓ^* satisfying

$$v'(\ell^*)\ell^* = 1 + \beta.$$

As a result, the planner faces

$$\begin{aligned} \max_{c_t^y,\, c_t^o} \quad & \mathbb{E}_0 \sum_{t=0}^{\infty} [\log c_t^y + \log c_t^o] \\ \text{s.t.} \quad & c_t^y + c_t^o \leq A_t \ell^* - A_t g_t \end{aligned}$$

which is merely a static endowment allocation problem, with an obvious solution: $c_t^y = c_t^o = \frac{A_t}{2}(\ell^* - g_t)$.[2] Using the individual's first order condition, we find taxes as

$$\begin{aligned} v'(\ell^*)c_t^y &= A_t(1 - \tau_t^L) \\ \frac{1}{2}v'(\ell^*)(\ell^* - g_t) &= 1 - \tau_t^L \\ \frac{1+\beta}{2}\left(1 - \frac{g_t}{\ell^*}\right) &= 1 - \tau_t^L \end{aligned}$$

[2] I've multiplied g_t by A_t for convenience; it is without loss of generality.

which is a function of only the current state.

To show that the distortionary cost is also a function only of the current state, consider the unaltered planner's problem:

$$\begin{aligned}
\max_{c_t^y,\, c_t^o} \quad & \mathbb{E}_0 \sum_{t=0}^{\infty} [\log c_t^y + \log c_t^o - v(\ell_t)] \\
\text{s.t.} \quad & c_t^y + c_t^o + A_t g_t \leq A_t \ell_t \\
& 1 + \beta - v'(\ell_t)\ell_t \geq 0
\end{aligned}$$

The first order condition with respect to ℓ_t is

$$A_t \psi_t = v'(\ell_t) + [v''(\ell_t)\ell_t + v'(\ell_t)]\mu_t$$

while the first order condition with respect to c_t^y is $\psi_t = 1/c_t^y$. Substituting in the values from above for ℓ_t and c_t^y and rearranging yields

$$\tilde{\mu}_t = \frac{\frac{2}{\ell^* - g_t} - v'(\ell^*)}{v''(\ell^*)\ell^* + v'(\ell^*)}$$

and so the distortionary cost of taxation depends only on the current state of the world and is not intertemporally linked in any way.

A.3 Proofs from Section 1.4

A.3.1 Proof of Proposition 1.4.1

This expression combines the first order condition for $z(s^{t+1})$ with the household's Euler equation for the asset paying off at history s^{t+1}.

A.3.2 Proof of Proposition 1.4.2

This expression combines the first order conditions with respect to c^y_{t+1} and c^o_{t+1} with the household's Euler equation for the asset paying off at history s^{t+1}.

A.3.3 Proof of Proposition 1.4.3

This combines the first order conditions with respect to c^y_t, c^o_{t+1}, and k_t, then employs the no-arbitrage condition between state-contingent assets and capital.

A.3.4 Proof of Corollary 1.4.4

This follows directly from differentiation of the utility function, rearranging, and employing the no-arbitrage condition.

A.3.5 Proof of Proposition 1.4.5

This expression combines the first order conditions for $z(s^{t+1})$ and $b(s^t)$ with the household's Euler equation for risk free debt issued at s^t.

A.3.6 Proof of Proposition 1.4.6

The first order condition with respect to b_t when debt limits do not bind is

$$u^y_{c,t}\mu^y_t = \beta R^f_t \mathbb{E}_t \mu^o_{t+1}.$$

The first order condition with respect to c^y_t when bequest constraints bind (and so

$\tilde{\xi}_t^\delta = 0$) is

$$\begin{aligned}\psi_t =& \tilde{w}_t + \mu_t^y \left[u_{c,t}^y + u_{cc,t}^y (c_t^y + k_t + b_t - z_t) + u_{c\ell,t}^y \ell_t\right] \\ & - \beta \mathbb{E}_t \left\{ \mu_{t+1}^o \left[\frac{u_{cc,t}^y}{\beta \mathbb{E}_t u_{c,t+1}^o} b_t + \frac{u_{cc,t}^y F_{K,t+1}}{\beta \mathbb{E}_t [F_{K,t+1} u_{c,t+1}^o]} k_t \right] \right\} \\ =& \tilde{w}_t \left\{ 1 + \tilde{\mu}_t \left[\eta_t - \sigma_t^y \frac{k_t + b_t}{c_t^y} \right] \right\} + \sigma_t^y \frac{b_t}{c_t^y} \beta R_t^f \mathbb{E}_t \mu_{t+1}^o + \sigma_t^y \frac{k_t}{c_t^y} \beta (1 - \tau_t^K) \mathbb{E}_t [F_{K,t+1} \mu_{t+1}^o] \\ =& \tilde{w}_t \left[1 + \tilde{\mu}_t \eta_t\right] + \sigma_t^y \frac{k_t}{c_t^y} \left\{ w_t \beta (1 - \tau_t^k) \mathbb{E}_t [F_{K,t+1} u_{c,t+1}^o \hat{\mu}_{t+1}] - \tilde{w}_t \tilde{\mu}_t \right\} \\ =& \tilde{w}_t \left[1 + \tilde{\mu}_t \eta_t\right] + \sigma_t^y \frac{k_t}{c_t^y} \left\{ \tilde{w}_t (1 - \tau_t^k) \mathbb{E}_t \left[F_{K,t+1} \frac{\beta u_{c,t+1}^o}{u_{c,t}^y} \hat{\mu}_{t+1} \right] - \tilde{w}_t \tilde{\mu}_t \right\} \\ =& \tilde{w}_t \left\{ 1 + \tilde{\mu}_t \eta_t + \frac{\sigma_t^y (1 - \tau_t^K)}{c_t^y} \left(\sum_{s_{t+1}} \left[q_{t+1} F_{K,t+1} k_t \hat{\mu}_{t+1}\right] - \frac{\tilde{\mu}_t k_t}{1 - \tau_t^K} \right) \right\}\end{aligned}$$

The first order condition with respect to c_t^o when bequest constraints bind (and so $\tilde{\xi}_t^\delta = 0$) is

$$\begin{aligned}\frac{\Delta}{\beta} \psi_t =& w_{t-1} u_{c,t}^o + \mu_t^o + \frac{u_{c,t-1}^y u_{cc,t}^o}{\beta (\mathbb{E}_{t-1} u_{c,t}^o)^2} b_{t-1} \mathbb{E}_{t-1} \mu_t^o + \frac{u_{c,t-1}^y F_{K,t} u_{cc,t}^o}{\beta (\mathbb{E}_{t-1} [F_{K,t} u_{c,t}^o])^2} k_{t-1} \mathbb{E}_{t-1} [\mu_t^o F_{K,t}] \\ & w_{t-1} u_{c,t}^o + \mu_t^o + R_{t-1}^f \frac{u_{cc,t}^o}{\mathbb{E}_{t-1} u_{c,t}^o} b_{t-1} \mathbb{E}_{t-1} \mu_t^o \\ & + (1 - \tau_{t-1}^K) \frac{F_{K,t} u_{cc,t}^o}{\mathbb{E}_{t-1} [F_{K,t} u_{c,t}^o]} k_{t-1} \mathbb{E}_{t-1} [\mu_t^o F_{K,t}]\end{aligned}$$

Now take expectations of both sides at date $t-1$:

$$\begin{aligned}\frac{\Delta}{\beta} \mathbb{E}_{t-1} \psi_t =& w_{t-1} \mathbb{E}_{t-1} u_{c,t}^o + \mathbb{E}_{t-1} \mu_t^o + R_{t-1}^f \frac{\mathbb{E}_{t-1} u_{cc,t}^o}{\mathbb{E}_{t-1} u_{c,t}^o} b_{t-1} \mathbb{E}_{t-1} \mu_{t-1}^o \\ & + (1 - \tau_{t-1}^K) \frac{\mathbb{E}_{t-1} [F_{K,t} u_{cc,t}^o]}{\mathbb{E}_{t-1} [F_{K,t} u_{c,t}^o]} k_{t-1} \mathbb{E}_{t-1} [\mu_t^o F_{K,t}]\end{aligned}$$

Using the first order condition for b_{t-1} and the definition of $\epsilon_{b_t}^{R_t^f}$ developed before leaves

$$\Delta \mathbb{E}_{t-1} \psi_t = \frac{\tilde{w}_{t-1}}{R_{t-1}^f} \left[1 + \tilde{\mu}_{t-1} \left(1 - \epsilon_{b_{t-1}}^{R_{t-1}^f} \right) \right] + (1 - \tau_{t-1}^K) \frac{\mathbb{E}_{t-1} [F_{K,t} u_{cc,t}^o]}{\mathbb{E}_{t-1} [F_{K,t} u_{c,t}^o]} k_{t-1} \mathbb{E}_{t-1} [\mu_t^o F_{K,t}]$$

Now differentiate

$$u^y_{c,t} = \beta(1-\tau^K_t)\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]$$

with respect to c^o_{t+1}, holding fixed $u^y_{c,t}$ and k_t, to obtain

$$0 = \beta(1-\tau^K_t)\mathbb{E}_t[F_{K,t+1}u^o_{cc,t+1}dc^o_{t+1}] - d\tau^K_t\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]$$

Considering a perturbation of c^o_{t+1} that involves only an increase in b_t, and therefore a rise in c^o_{t+1} by R^f_t across all states, yields

$$\frac{\partial\tau^K_t}{\partial b_t} = (1-\tau^K_t)R^f_t\frac{\mathbb{E}_t[F_{K,t+1}u^o_{cc,t+1}]}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]}.$$

Substituting this back in yields

$$\Delta\mathbb{E}_{t-1}\psi_t = \frac{\tilde{w}_{t-1}}{R^f_{t-1}}\left[1+\tilde{\mu}_{t-1}\left(1-\epsilon^{R^f_{t-1}}_{b_{t-1}}\right)\right] + \frac{1}{R^f_{t-1}}\frac{\partial\tau^K_t}{\partial b_t}k_{t-1}\mathbb{E}_{t-1}[F_{K,t}\mu^o_t].$$

Replacing $\mathbb{E}_{t-1}[F_{K,t}\mu^o_t]$ as before yields

$$\Delta\mathbb{E}_{t-1}\psi_t = \frac{\tilde{w}_{t-1}}{R^f_{t-1}}\left[1+\tilde{\mu}_{t-1}\left(1-\epsilon^{R^f_{t-1}}_{b_{t-1}}\right)\right] + \frac{1}{R^f_{t-1}}\frac{\partial\tau^K_t}{\partial b_t}k_{t-1}\tilde{w}_{t-1}\sum_{s_t}[q_tF_{K,t}\hat{\mu}_t]$$

Combining this with the expectation at date $t-1$ of the first order condition with respect to young consumption at date t, and then advancing the time index yields the desired result.

A.3.7 Proof of Proposition 1.4.7

In addition to the previous first order conditions, take the first order condition with respect to k_t:

$$\begin{aligned}\psi_t =& \Delta\mathbb{E}_t[F_{K,t+1}\psi_{t+1}] + \mu^y_t u^y_{c,t} - u^y_{c,t}\frac{\mathbb{E}_t[F_{K,t+1}\mu^o_{t+1} + F_{KK,t+1}k_t\mu^o_{t+1}]}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]}\\ &+ u^y_{c,t}\frac{\mathbb{E}_t[k_tF_{K,t+1}\mu^o_{t+1}]}{(\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}])^2}\mathbb{E}_t[F_{KK,t+1}u^o_{c,t+1}]\end{aligned}$$

The last two terms will clearly give rise to the "hedging" term of Farhi (2010) and the quasilinear example, so I'll start by simplifying just those. The last term reduces to

$$\beta w_t(1-\tau_t^K)\frac{\mathbb{E}_t[F_{KK,t+1}u^o_{c,t+1}]}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]}\left\{k_t\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]\mathbb{E}_t\hat{\mu}_{t+1}+\mathrm{Cov}_t[\hat{\mu}_{t+1},k_tF_{K,t+1}u^o_{c,t+1}]\right\}$$

The penultimate term reduces to

$$\begin{aligned}&-\beta w_t(1-\tau_t^K)\left\{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}\hat{\mu}_{t+1}]+\mathbb{E}_t[F_{KK,t+1}k_tu^o_{c,t+1}]\mathbb{E}_t\hat{\mu}_{t+1}\right.\\&\left.+\mathrm{Cov}_t[\hat{\mu}_{t+1},k_tF_{KK,t+1}u^o_{c,t+1}]\right\}\end{aligned}$$

Summing them yields

$$\begin{aligned}\beta w_t(1-\tau_t^K)&\left\{\frac{\mathbb{E}_t[F_{KK,t+1}u^o_{c,t+1}]}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]}\mathrm{Cov}_t[\hat{\mu}_{t+1},k_tF_{K,t+1}u^o_{c,t+1}]-\mathrm{Cov}_t[\hat{\mu}_{t+1},k_tF_{KK,t+1}u^o_{c,t+1}]\right.\\&\left.-\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}\hat{\mu}_{t+1}]\right\}.\end{aligned}$$

I summarize the first two of these terms as $-\beta w_t(1-\tau_t^K)H.T.$ (hedging terms) and simplify the first order condition to

$$\psi_t=\Delta\mathbb{E}_t[F_{K,t+1}\psi_{t+1}]+\mu_t^yu^y_{c,t}-\beta w_t(1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}\hat{\mu}_{t+1}]-\beta w_t(1-\tau_t^K)H.T.$$

Now substitute in for ψ_t using the first order condition with respect to c_t^y and for ψ_{t+1} using the first order condition with respect to c^o_{t+1}:

$$\begin{aligned}&w_tu^y_{c,t}\left\{1+\tilde{\mu}_t\eta_t+\frac{\sigma_t^y(1-\tau_t^K)}{c_t^y}\left(\sum_{s_{t+1}}[q_{t+1}F_{K,t+1}k_t\hat{\mu}_{t+1}]-\frac{\tilde{\mu}_tk_t}{1-\tau_t^K}\right)\right\}\\=&\beta\mathbb{E}_t\left\{F_{K,t+1}\left[w_tu^o_{c,t+1}+\mu^o_{t+1}+R_t^f\frac{u^o_{cc,t+1}}{\mathbb{E}_tu^o_{c,t+1}}b_t\mathbb{E}_t\mu^o_{t+1}\right.\right.\\&\left.\left.+(1-\tau_t^K)\frac{F_{K,t+1}u^o_{cc,t+1}}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]}k_t\mathbb{E}_t[\mu^o_{t+1}F_{K,t+1}]\right]\right\}\\&+\mu_t^yu^y_{c,t}-\beta w_t(1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}\hat{\mu}_{t+1}]-\beta w_t(1-\tau_t^K)H.T.\end{aligned}$$

Next, subtract $\mu_t^y u_{c,t}^y$ from both sides and expand q_{t+1} and $\mu_{t+1}^o = w_t u_{c,t+1}^o \hat{\mu}_{t+1}$:

$$
\begin{aligned}
& w_t\beta(1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o][1+\tilde{\mu}_t(\eta_t-1)] \\
& + w_t(1-\tau_t^K)\sigma_t^y\frac{k_t}{c_t^y}\mathbb{E}_t[\beta u_{c,t+1}^o F_{K,t+1}\hat{\mu}_{t+1}] \\
& - w_t\beta(1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o]\sigma_t^y\tilde{\mu}_t\frac{k_t}{c_t^y} \\
=& \beta w_t\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o] + \beta w_t\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o\hat{\mu}_{t+1}] + \beta w_t R_t^f\tilde{\mu}_t b_t\mathbb{E}_t[F_{K,t+1}u_{cc,t+1}^o] \\
& + \beta w_t(1-\tau_t^K)\frac{\mathbb{E}_t[(F_{K,t+1})^2u_{cc,t+1}^o]}{\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o]}k_t\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o\hat{\mu}_{t+1}] \\
& - \beta w_t(1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o\hat{\mu}_{t+1}] - \beta w_t(1-\tau_t^K)H.T.
\end{aligned}
$$

Divide through by βw_t and subtract $\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o]$ from both sides:

$$
\begin{aligned}
& -\tau_t^K\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o] + (1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o]\tilde{\mu}_t(\eta_t-1) \\
& + (1-\tau_t^K)\sigma_t^y\frac{k_t}{c_t^y}\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o(\hat{\mu}_{t+1}-\tilde{\mu}_t)] \\
=& \mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o\hat{\mu}_{t+1}] + R_t^f\tilde{\mu}_t b_t\mathbb{E}_t[F_{K,t+1}u_{cc,t+1}^o] \\
& + (1-\tau_t^K)\frac{\mathbb{E}_t[(F_{K,t+1})^2u_{cc,t+1}^o]}{\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o]}k_t\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o\hat{\mu}_{t+1}] \\
& - (1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u_{c,t+1}^o\hat{\mu}_{t+1}] - (1-\tau_t^K)H.T.
\end{aligned}
$$

Next define two concepts analogous to those defined in the last proof. Start with the household Euler equation with respect to risk free bonds:

$$u_{c,t}^y = \beta R_t^f\mathbb{E}_t[u_{c,t+1}^o]$$

Consider the differential form holding c_t^y constant:

$$0 = R_t^f\mathbb{E}_t[u_{cc,t+1}^o dc_{t+1}^o] + dR_t^f\mathbb{E}_t[u_{c,t+1}^o]$$

Now suppose that $dc_{t+1}^o = (1-\tau_t^K)F_{K,t+1}dk_t$, as if accumulation of k_t were increased.

Then write

$$\frac{\partial R_t^f}{\partial k_t} = -\frac{R_t^f(1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u^o_{cc,t+1}]}{\mathbb{E}_t u^o_{c,t+1}}$$

which immediately yields

$$\epsilon_{k_t}^{R_t^f} = -(1-\tau_t^K)\frac{\mathbb{E}_t[F_{K,t+1}u^o_{cc,t+1}]}{\mathbb{E}_t u^o_{c,t+1}}k_t$$

Likewise consider the household Euler equation with respect to capital in differential form, holding c_t^y constant:

$$0 = (1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u^o_{cc,t+1}dc^o_{t+1}] - d\tau_t^K\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]$$

Supposing again that $dc^o_{t+1} = (1-\tau_t^K)F_{K,t+1}dk_t$ we have

$$\frac{\partial \tau_t^K}{\partial k_t} = (1-\tau_t^K)^2\frac{\mathbb{E}_t[(F_{K,t+1})^2u^o_{cc,t+1}]}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]}$$

which yields

$$\epsilon_{k_t}^{1-\tau_t^K} = -(1-\tau^K)\frac{\mathbb{E}_t[(F_{K,t+1})^2u^o_{cc,t+1}]}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]}k_t$$

Substituting in these two expressions into the main equation gives

$$\begin{aligned}
&-\tau_t^K\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}] + (1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]\tilde{\mu}_t(\eta_t-1)\\
&+(1-\tau_t^K)\sigma_t^y\frac{k_t}{c_t^y}\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}(\hat{\mu}_{t+1}-\tilde{\mu}_t)]\\
=&\tau_t^K\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}\hat{\mu}_{t+1}] - \frac{R_t^f}{1-\tau_t^K}\tilde{\mu}_t\frac{b_t}{k_t}\mathbb{E}_t[u^o_{c,t+1}]\epsilon_{k_t}^{R_t^f}\\
&-\epsilon_{k_t}^{1-\tau_t^K}\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}\hat{\mu}_{t+1}] - (1-\tau_t^K)H.T.
\end{aligned}$$

Note that $R_t^f\mathbb{E}_t[u^o_{c,t+1}] = (1-\tau_t^K)\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]$ and combine terms:

$$\frac{\tau_t^K}{1-\tau_t^K} = \frac{\left\{\begin{array}{c}\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}]\tilde{\mu}_t(\eta_t-1) + \sigma_t^y\frac{k_t}{c_t^y}\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}(\hat{\mu}_{t+1}-\tilde{\mu}_t)]\\ +\frac{1}{1-\tau_t^K}\mathbb{E}_t\left\{F_{K,t+1}u^o_{c,t+1}\left[\epsilon_{k_t}^{R_t^f}\frac{b_t}{k_t}\tilde{\mu}_t + \epsilon_{k_t}^{1-\tau_t^K}\hat{\mu}_{t+1}\right]\right\} + H.T.\end{array}\right\}}{\mathbb{E}_t[F_{K,t+1}u^o_{c,t+1}(1+\hat{\mu}_{t+1})]}$$

Reordering gives the desired result.

A.4 Further Details about Numerical Simulations

A.4.1 Recursive Formulations of Problems 1.2.1 and 1.2.2

In all cases, the initial period is different, and will be treated separately. I focus only on an "OLG" case with no bequest motive and a zero bequest limit, or a "Ramsey" case with perfect bequest motives ($\delta = \Delta = \beta$) and no bequest limits, since these are simpler to specify and are the only cases I simulate.

I allow for trend growth a in the following manner. For $x \in \{A, g, c^y, c\}$, I redefine $x_t = \exp(at)x$. I redefine $c_t^o = \exp(a(t-1))c_t^o$ and $k_t = \exp(a(t+1))k_t$. Then, if $u(\bullet)$ is log-log and $W(\bullet)$ is CARA, we can substitute $\ell = \ell^*$ as the result of the implementability condition since labor is constant[3], and we have

Problem A.4.1 (OLG Complete Markets, Log-Log) *Define*

$$\begin{aligned}
V(\theta, s_-) = \max_{c^y, \boldsymbol{c}_s^o, k'} \quad & W(u(c^y, \ell^*) + \beta\mathbb{E}[u(c_s^o, 0)|s_-]) \\
& + \exp(-\zeta(1-\gamma)a)\Delta\mathbb{E}\left[V\left(F(k', \ell^*; A_s) - \frac{c_s^o}{\exp(a)} - g_s, s\right)|s_-\right] \\
\text{s.t.} \quad & c^y + \exp(a)k' \leq \theta
\end{aligned}$$

[3] assuming it binds, which is does in all of my simulations

Then the planner solves

$$\begin{aligned}
\max_{c^o_-,\, c^y,\, \boldsymbol{c}^o_s,\, k'} \quad & \exp(\zeta(1-\gamma)a)\Delta^{-1}W(u(c^y_-,\ell_-)+\beta u(c^o_-,0)) \\
& +W(u(c^y,\ell)+\beta\mathbb{E}[u(c^o_s,0)|s_-]) \\
& +\exp(-\zeta(1-\gamma)a)\Delta\mathbb{E}\left[V\left(F(k',\ell^*;A_s)-\frac{c^o_s}{\exp(a)}-g_s,s\right)|s_-\right] \\
\text{s.t.} \quad & c^y+\frac{c^o_-}{\exp(a)}+g_{s_-}+\exp(a)k' \leq F(k,\ell^*;A_{s_-}) \\
& c^o_- \geq F_K(k,\ell^*;A_{s_-})k+b_0
\end{aligned}$$

If $u(\bullet)$ is isoelastic and $W(\bullet)$ is CRRA, then we have

Problem A.4.2 (OLG Complete Markets, Isoelastic) *Define*

$$\begin{aligned}
V(k,c^o,s_-) = \max_{c^y,\,\ell,\,\boldsymbol{c}^o_s,\,k'} \quad & W(u(c^y,\ell)+\beta\mathbb{E}[u(c^o_s,0)|s_-])+ \\
& \exp((1-\zeta)(1-\sigma)a)\Delta\mathbb{E}\left[V(k',c^o_s,s)\right] \\
\text{s.t.} \quad & c^y+\frac{c^o}{\exp(a)}+g_{s_-}+\exp(a)k' \leq F(k,\ell;A_{s_-}) \\
& u_c(c^y,\ell)c^y+u_\ell(c^y,\ell)\ell+\beta\mathbb{E}[u_c(c^o_s,0)c^o_s|s_-] \geq 0
\end{aligned}$$

Then the planner solves

$$\begin{aligned}
\max_{c^o,\,c^y,\,\ell,\,\boldsymbol{c}^o_s,\,k'} \quad & \exp((1-\zeta)(1-\sigma)a)\Delta^{-1}W(u(c^y_-,\ell_-)+\beta u(c^o_-,0)) \\
& +W(u(c^y,\ell)+\beta\mathbb{E}[u(c^o_s,0)|s_-])+\exp((1-\zeta)(1-\sigma)a)\Delta\mathbb{E}V(k',c^o_s,s) \\
\text{s.t.} \quad & c^y+\frac{c^o_-}{\exp(a)}+g(s_-)+\exp(a)k' \leq F(k,\ell;A_{s_-}) \\
& c^o_- \geq F_K(k,\ell;A)k+b_0 \\
& u_c(c^y,\ell)c^y+u_\ell(c^y,\ell)\ell+\beta\mathbb{E}[u_c(c^o_s,0)c^o_s|s_-] \geq 0
\end{aligned}$$

For Ramsey models, I assume there is no trend growth[4] and that utility is separable between labor and consumption. This ensures that $c^y(s^t) = c^o(s^t) \equiv c(s^t0)$.

Problem A.4.3 (Ramsey Complete Markets) *Define*

$$
\begin{aligned}
V(k,\theta,s_-) \quad \max_{c,\ell,k',\boldsymbol{\theta}'_s} \quad & u(c,\ell) + u(c,0) + \Delta\mathbb{E}[V(k',\theta'_s,s)|s_-] \\
\text{s.t.} \quad & 2c + g_{s_-} + k' \leq F(k,\ell;A_{s_-}) \\
& u_c(c,\ell)c + u_\ell(c,\ell)\ell + u_c(c,0)c + \beta\mathbb{E}[\theta'_s|s_-] \geq \theta
\end{aligned}
$$

Then the planner solves

$$
\begin{aligned}
\max_{c,\ell,k',\boldsymbol{\theta}'_s} \quad & u(c,\ell) + u(c,0) + \Delta\mathbb{E}[V(k',\theta'_s,s)|s_-] \\
\text{s.t.} \quad & 2c + g_{s_-} + k' \leq F(k,\ell;A_{s_-}) \\
& u_c(c,\ell)c + u_\ell(c,\ell)\ell + u_c(c^o,0) + \beta\mathbb{E}[\theta'_s|s_-] \geq u_c(c,\ell)[F_K(k,\ell;A_{s_-})k + b_0] \\
& u_c(c^y,\ell) = u_c(c^o,0).
\end{aligned}
$$

If utility is log-log, the incomplete markets model simplifies substantially—especially the old implementability condition.

---

[4] All models with trend growth that I simulate have an obvious first best implementation in a Ramsey model.

Problem A.4.4 (OLG Incomplete Markets, Log-Log) *Define*

$$V(\theta, l, s_-) = \max_{c^y, \boldsymbol{c}^o_s, k', b', \boldsymbol{\theta}'_s, \boldsymbol{\ell}'_s} \quad W(u(c^y, \ell) + \beta \mathbb{E}[u(c^o_s, 0)|s_-]) + \exp(-\zeta(1-\gamma)a)\Delta$$

$$\bullet\, \mathbb{E}\left[V\left(F(k', \ell_s; A_s) - \frac{c^o_s}{\exp(a)} - g_s, \ell'_s, s\right)|s_-\right]$$

s.t.

$$c^y + \exp(a)k' \leq \theta$$

$$u_c(c^y, \ell)(c^y + b' + \exp(a)k') + u_\ell(c^y, \ell)\ell \geq 0$$

$$c^y \geq \frac{1/c^o_s}{\beta \mathbb{E}[1/c^o_s|s_-]} b' + \frac{F_K(k', \ell'_s; A_s)/c^o_s}{\beta \mathbb{E}[F_K(k', \ell'_s; A_s)/c^o_s|s_-]} \exp(a)k'$$

$$\underline{b} \leq b' \leq \overline{b}$$

Then the planner solves

$$\max_{c^o, \theta', \ell'} \quad W(u(c^y_-, \ell_-)$$

$$+\beta u(c^o, 0)) + \exp(-\zeta(1-\gamma)a)\Delta V\left(F(k_-, \ell'; A_{s_-}) - \frac{c^o}{\exp(a)} - g_{s_-}, \ell', s_-\right)$$

$$\text{s.t.} \qquad c^o \geq R^f_- b_- + (1-\tau^K_-)F_K(k_-, \ell'; A_{s_-})k_-$$

In principle, this formulation could be used almost identically for other IESs. Unfortunately, in practice, finding a rectangular state space that encloses the ergodic set is impossible. Thus, I change the recursive formulation for isoelastic utility to one that is defined *ex interim*:

Problem A.4.5 (OLG Incomplete Markets, Isoelastic) *Define*

$$
\begin{aligned}
V(k,\tilde{a},\tilde{b},v,s_-) = \max_{\boldsymbol{c}^y_s,\boldsymbol{c}^o_s,\boldsymbol{\ell}_s,\boldsymbol{k}'_s,\tilde{\boldsymbol{a}}'_s,\tilde{\boldsymbol{b}}'_s} \quad & W(v+\beta\mathbb{E}[u(c^o_s,0)|s_-]) \\
& + \exp((1-\zeta)(1-\sigma)a)\Delta\mathbb{E}\left[V(k'_s,\tilde{a}'_s,\tilde{b}'_s,u(c^y_s,\ell_s),s)|s_-\right] \\
& \text{s.t.} \\
& c^y_s + \frac{c^o_s}{\exp(a)} + g_s + \exp(a)k'_s \leq F(k,\ell_s;A_s) && \forall s\in\mathcal{S} \\
& u_c(c^y_s,\ell_s)c^y_s + u_\ell(c^y_s,\ell_s)\ell_s + \tilde{a}'_s \geq 0 && \forall s\in\mathcal{S} \\
& c^o_s \geq \frac{\tilde{b}}{\beta\mathbb{E}[u_c(c^o_s,0)|s_-]} + \frac{F_K(k,\ell_s;A_s)(\tilde{a}-\tilde{b})}{\beta\mathbb{E}[F_K(k,\ell_s;A_s)u_c(c^o_s,0)|s_-]} && \forall s\in\mathcal{S} \\
& \underline{b} \leq \frac{\tilde{b}'_s}{u_c(c^y_s,\ell_s)} \leq \bar{b} && \forall s\in\mathcal{S} \\
& \tilde{a}'_s - \tilde{b}'_s = \exp(a)k'_s u_c(c^y_s,\ell_s) && \forall s\in\mathcal{S}
\end{aligned}
$$

Then the planner solves

$$
\begin{aligned}
\max_{c^y,c^o,\ell,k',\tilde{a}',\tilde{b}'} \quad & W(v+\beta u(c^o,0)) + \exp((1-\zeta)(1-\sigma)a)\Delta V(k',\tilde{a}',\tilde{b}',u(c^y,\ell),s) \\
\text{s.t.} \quad & c^y + \frac{c^o}{\exp(a)} + g_{s_-} + \exp(a)k' \leq F(k,\ell;A_{s_-}) \\
& u_c(c^y,\ell)c^y + u_\ell(c^y,\ell)\ell + \tilde{a}' \geq 0 \\
& c^o \geq R^f_- b_- + (1-\tau^K_-)F_K(k,\ell,A_{s_-})k \\
& \underline{b} \leq \frac{\tilde{b}'}{u_c(c,\ell)} \leq \bar{b}
\end{aligned}
$$

Problem A.4.6 (Ramsey Incomplete Markets) *Define*

$$
\begin{aligned}
V(k,\tilde{a},\tilde{b},s_-) = \max_{\boldsymbol{c}_s,\boldsymbol{\ell}_s,\boldsymbol{k}'_s,\tilde{\boldsymbol{a}}'_s,\tilde{\boldsymbol{b}}'_s,} \quad & \mathbb{E}\left\{[u(c_s,\ell_s)+u(c_s,0)+\Delta V(k'_s,\tilde{a}'_s,\tilde{b}'_s,s)]|s_-\right\} \\
\text{s.t.} \quad & \\
2c_s + g_s + k'_s \leq & F(k,\ell_s;A_s) \qquad \forall s\in\mathcal{S} \\
\left\{\begin{array}{c} 2u_c(c_s,\ell_s)c_s \\ +u_\ell(c_s,\ell_s)\ell_s + \tilde{a}'_s \end{array}\right\} \geq & \frac{\tilde{b}u_c(c_s,\ell_s)}{\beta\mathbb{E}[u_c(c_s,\ell_s)|s_-]} \\
& + \frac{F_K(k,\ell_s;A_s)u_c(c_s,\ell_s)(\tilde{a}-\tilde{b})}{\beta\mathbb{E}[F_K(k,\ell_s;A_s)u_c(c_s,\ell_s)|s_-]} \qquad \forall s\in\mathcal{S} \\
\underline{b} \leq \frac{\tilde{b}'_s}{u_c(c_s,\ell_s)} \leq & \bar{b} \qquad \forall s \\
\tilde{a}'_s - \tilde{b}'_s = & k'_s u_c(c^y_s,\ell_s) \qquad \forall s\in\mathcal{S}
\end{aligned}
$$

Then the planner solves

$$
\begin{aligned}
\max_{c,\ell,k',\tilde{a}',\tilde{b}'} \quad & u(c,\ell)+u(c,0)+\Delta V(k',\tilde{a}',\tilde{b}',s_-) \\
\text{s.t.} \quad & \\
2c + g_{s_-} + k' \leq & F(k,\ell;A_{s_-}) \\
2u_c(c,\ell)c + u_\ell(c,\ell)\ell + \tilde{a}' \geq & u_c(c,\ell)[R^f_- b_- + (1-\tau^K_-)F_K(k,\ell;A_{s_-})k] \\
\underline{b} \leq \frac{\tilde{b}'}{u_c(c,\ell)} \leq & \bar{b}
\end{aligned}
$$

A.4.2 Further Details on Solution Method

As stated in the main text, my solution method was collocation using Chebyshev polynomials on a sparse grid; in the footnotes, I list a few references that explain the method in extensive detail, along with the reasons why various choices are made. Here, I merely give a brief explanation of the method in case the reader is unfamiliar. It differs from standard value function iteration in three major respects—the approximation function to the value function, the spacing of collocation points on the grid, and how the procedure scales with the number of state space dimensions.

In "standard" value function iteration, the value function is approximated using splines between equally spaced points. The simplest example would be linear splines, in which the value function between grid points is linearly interpolated using the two nearest points[5] on the grid; cubic splines, which require four points, are quite commonly used as well. By nature of spline interpolation, the interpolation function is local—it is different in different regions of the state space.

In contrast, the method I use approximates the value function with the same weighted sum of functions across the entire state space. For simplicity, assume that this space is $[0, 1]$, though scaling and translating it is a trivial procedure. Specifically, an Nth degree approximation in a single dimension is as follows:

$$V(x) \approx \sum_{n=1}^{N} \theta_n T_n(x)$$

where $\boldsymbol{\theta}$ is some set of coefficients to be solved for and $T_n(x)$ is the nth Chebyshev polynomial of the first kind, defined recursively by $T_1(x) = 1$, $T_2(x) = x$, $T_{n+1}(x) = 2x$, $T_n(x) - T_{n-1}(x)$. This approximation method exploits the fact that most value functions encountered in economics—such as the ones I work with in the present paper—are globally smooth, in order to vastly reduce the number of necessary grid points.

Instead of spline approximation, which suggests using collocation points—the points at which to enforce that the Bellman equation holds exactly—on the boundaries of the splines, Chebyshev approximation suggests using Chebyshev nodes (either zeros or extrema of the Chebyshev polynomials) as collocation points. In accordance with

[5] in each dimension

Krueger and Kubler (2004), I choose the extrema:

$$x_n = -\cos\left(\frac{n-1}{N-1}\right) \quad n = 1, 2, ..., N$$

The procedure in one dimension is then quite similar to value function iteration—to progressively better approximate the value function. In the case of Chebyshev polynomials, the equation that enforces that the Bellman equation holds perfectly at the collocation points is $\Phi\boldsymbol{\theta} = \boldsymbol{v}(\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ is the vector of coefficients; $\boldsymbol{v}(\boldsymbol{\theta})$ is the vector of results of the Bellman maximization problems; with typical element $v_j(\boldsymbol{\theta})$ the result of the maximization problem at point x_j; and Φ is the collocation matrix with typical element $\phi_{ij} = T_j(x_i)$, which is constant. The update procedure, suggested by Miranda and Fackler (1997), is then $\boldsymbol{\theta} := \Phi^{-1}\boldsymbol{v}(\boldsymbol{\theta})$, repeated until convergence, measured according to either the Bellman equation error at the collocation points, or the change in $\boldsymbol{\theta}$.

The final issue to consider is how to scale this procedure to multiple dimensions; I follow Judd et al. (2014).[6] The most straightforward way would be to use a tensor product of the T_n and x_n—that is, take the Cartesian product of the x_n, and take the Cartesian product, multiplied together, of the T_n. However, this procedure scales exponentially—there are N^d points and coefficients for d dimensions—making it a poor choice for larger-dimensional problems. Instead, we choose a grid which is denser near the steady state (the center of the state space) and the edges of the state space and sparser in between, in a way that scales polynomially in the number of dimensions d.

[6] In fact, the reason I originally learned and implemented this method is that I intended to solve problems featuring many dimensions—additional generations, and additional types within each generation. I have since decided to defer such investigations to future papers, but I hope this investment will pay dividends at that time.

Define a sequence of sets in a single dimension as follows. Let

$$S_i \equiv \bigcup_{n=1}^{m(i)} \left[-\cos\left(\frac{n-1}{m(i)-1} \right) \right],$$

where $m(i) \equiv 2^{i-1} + 1$ for $i \geq 2$, and $m(1) \equiv 1$. Then the set of points in the grid of dimension d using approximation level μ is

$$\Theta_\mu \equiv \bigcup_{\boldsymbol{i} | d \leq \sum_{j=1}^{d} i_j \leq d+\mu} [S_{i_1} \otimes S_{i_2} \otimes ... \otimes S_{i_d}]$$

Then choose the corresponding set of Chebyshev polynomials to include in the value function approximation.

A.5 Additional Simulations

A.5.1 Stochastic Government Expenditure

Complete Markets. Figures A.1 and A.2 respectively show the behavior of several variables across state transitions, and the autocorrelation structure of labor taxes, if the IES is 1/4 and markets are complete. The concave planner has $\zeta = -8$. Likewise, Figures A.3 and A.4 respectively show the behavior of several variables across state transitions, and the autocorrelation structure of labor taxes, if the IES is 4 and markets are complete. The concave planner has $\zeta = 8$. The results are qualitatively extremely similar to each other and the log-log case in the main text. The primary difference is that the planner chooses tax rates that depend more strongly on the current state and less on the previous state when the IES is low; this reflects the strong lack of Ricardian equivalence and incentive to run a balanced budget.

It is worth noting that comparing planners in different models is not straightforward; one must take a stand on whether the planner actually cares about inequality in cardinal

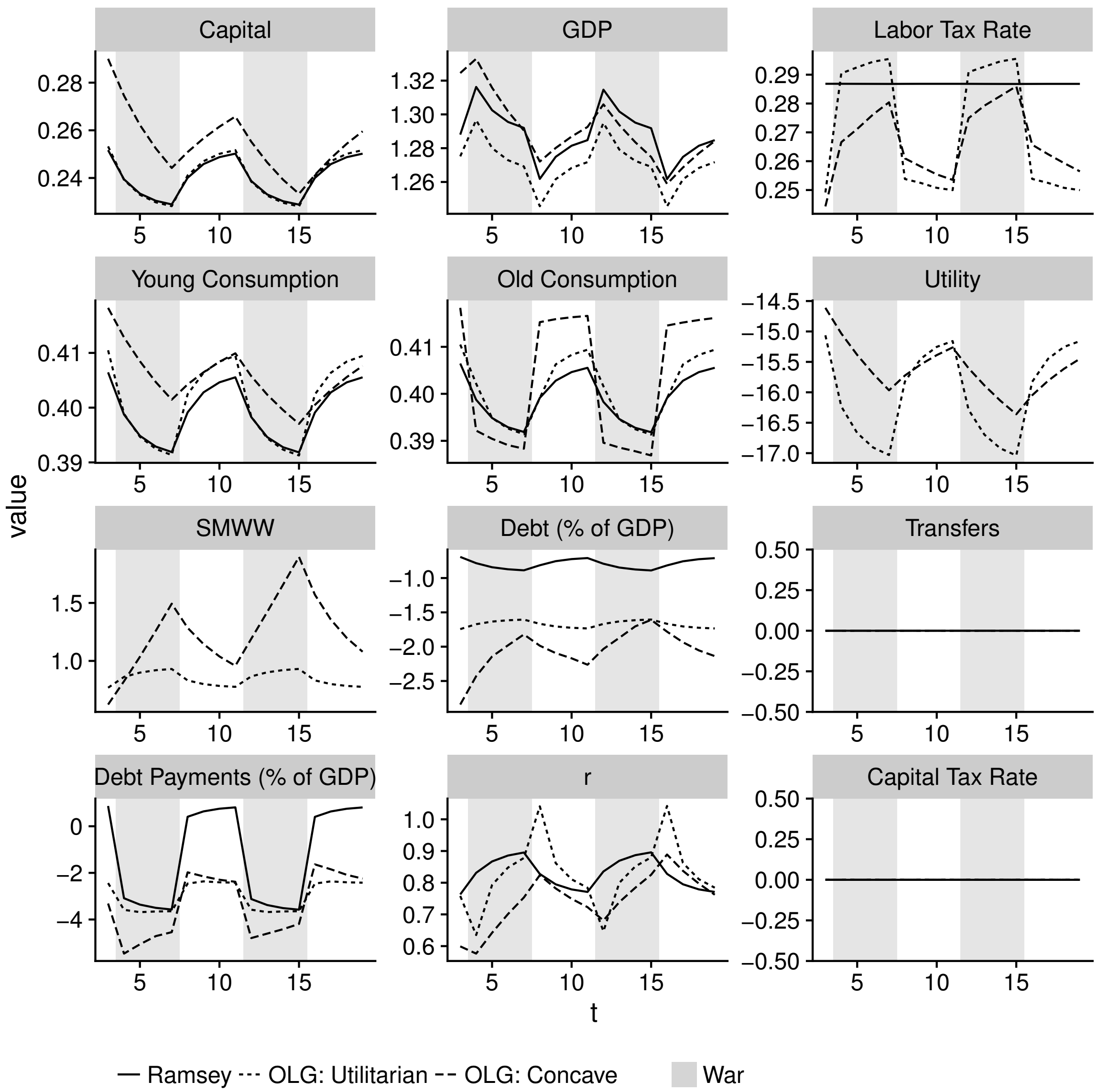


Figure A.1: This figure captures the salient features of state transitions in the complete markets model of stochastic government expenditure when utility is isoelastic with $\sigma = 4$. Shading marks periods of high government spending ("wars"). The concave planner has $\zeta = -8$. The model was first allowed to run for 100 periods in state 1 to arrive at a steady state. The Ramsey model was started with debt equal to -1.7% of nonstochastic steady state GDP, which corresponds to the level of government debt in the OLG model after 100 periods of low government spending; initial conditions are irrelevant for the OLG models.

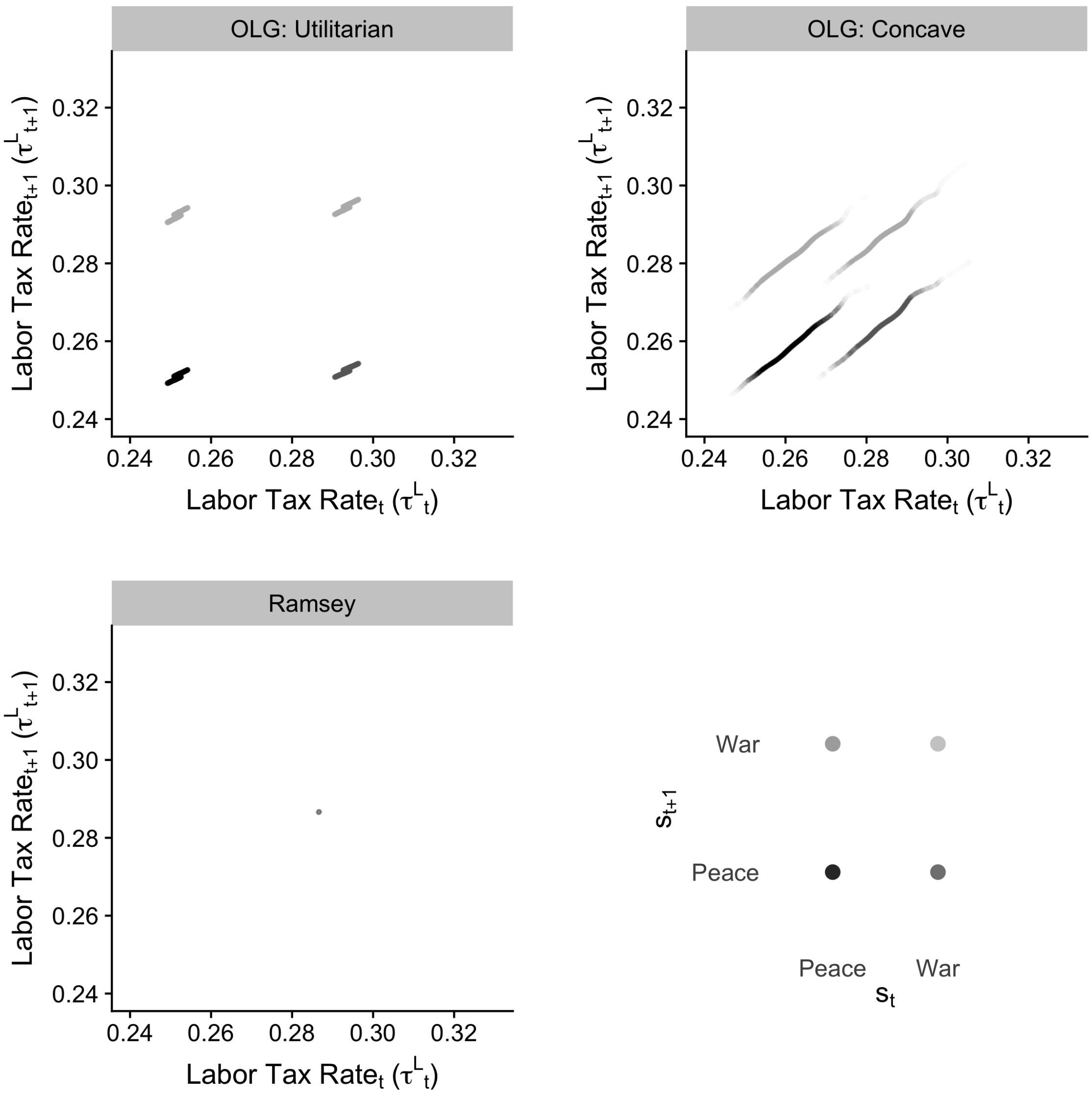


Figure A.2: This figure presents the results of a random simulation of the complete markets model of stochastic government expenditure when utility is isoelastic with $\sigma = 4$. The model is simulated over 10000 periods (after a 100-period startup period that is excluded), and each point represents one period. The concave planner has $\zeta = -8$. The labor tax rate at date t is on the horizontal axis, and the labor tax rate at date $t+1$ is on the vertical axis. The different shades of gray represent different sequences of shocks received at date t (s_t) and date $t+1$ (s_{t+1}). The Ramsey model was started with debt equal to -1.7% of nonstochastic steady state GDP, which corresponds to the level of government debt in the OLG model after 100 periods of low government spending; initial conditions are irrelevant for the OLG models.

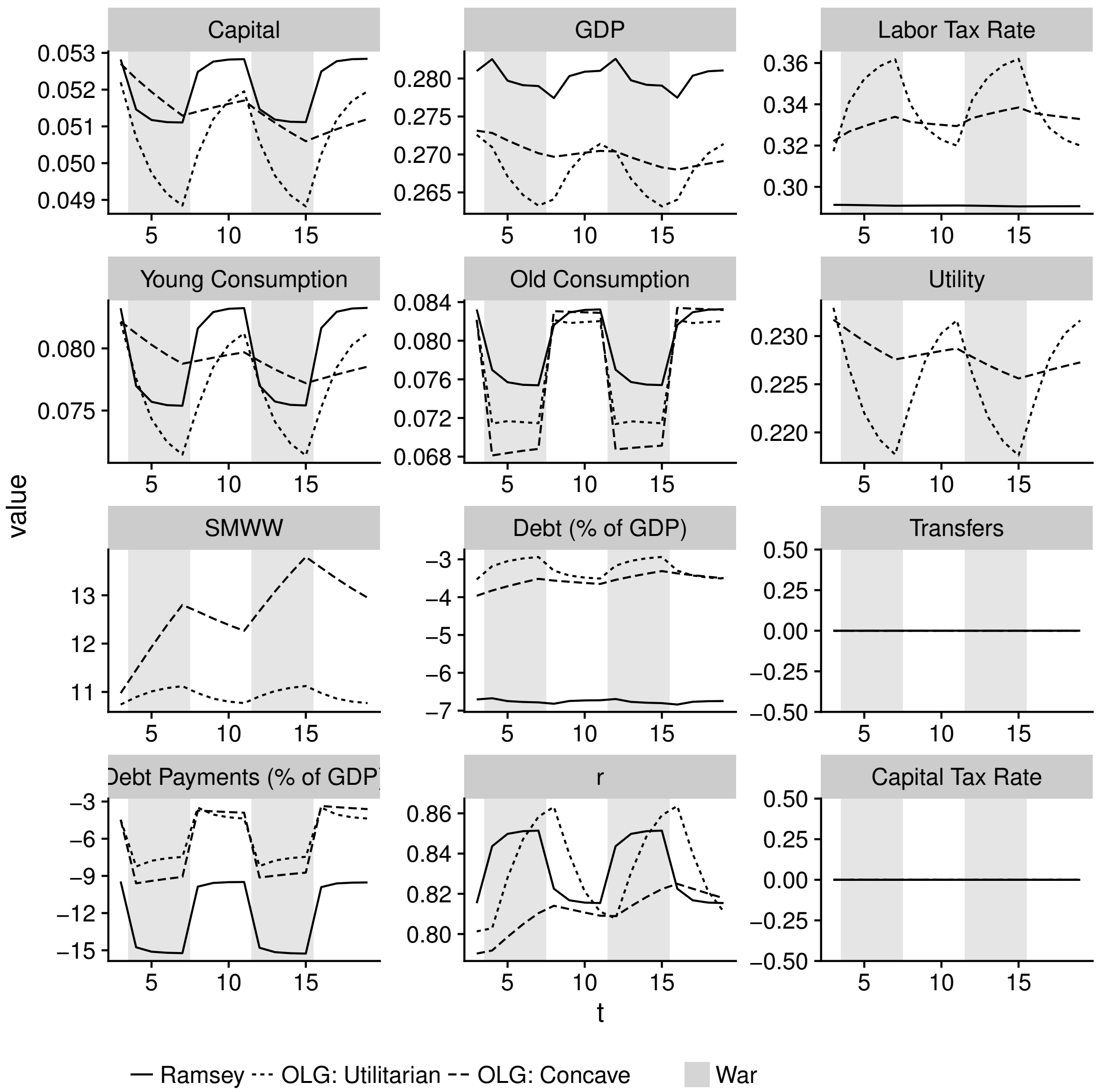


Figure A.3: This figure captures the salient features of state transitions in the complete markets model of stochastic government expenditure when utility is isoelastic with $\sigma = 1/4$. Shading marks periods of high government spending ("wars"). The concave planner has $\zeta = 8$. The model was first allowed to run for 100 periods in state 1 to arrive at a steady state. The Ramsey model was started with debt equal to -3.5% of nonstochastic steady state GDP, which corresponds to the level of government debt in the OLG model after 100 periods of low government spending; initial conditions are irrelevant for the OLG models.

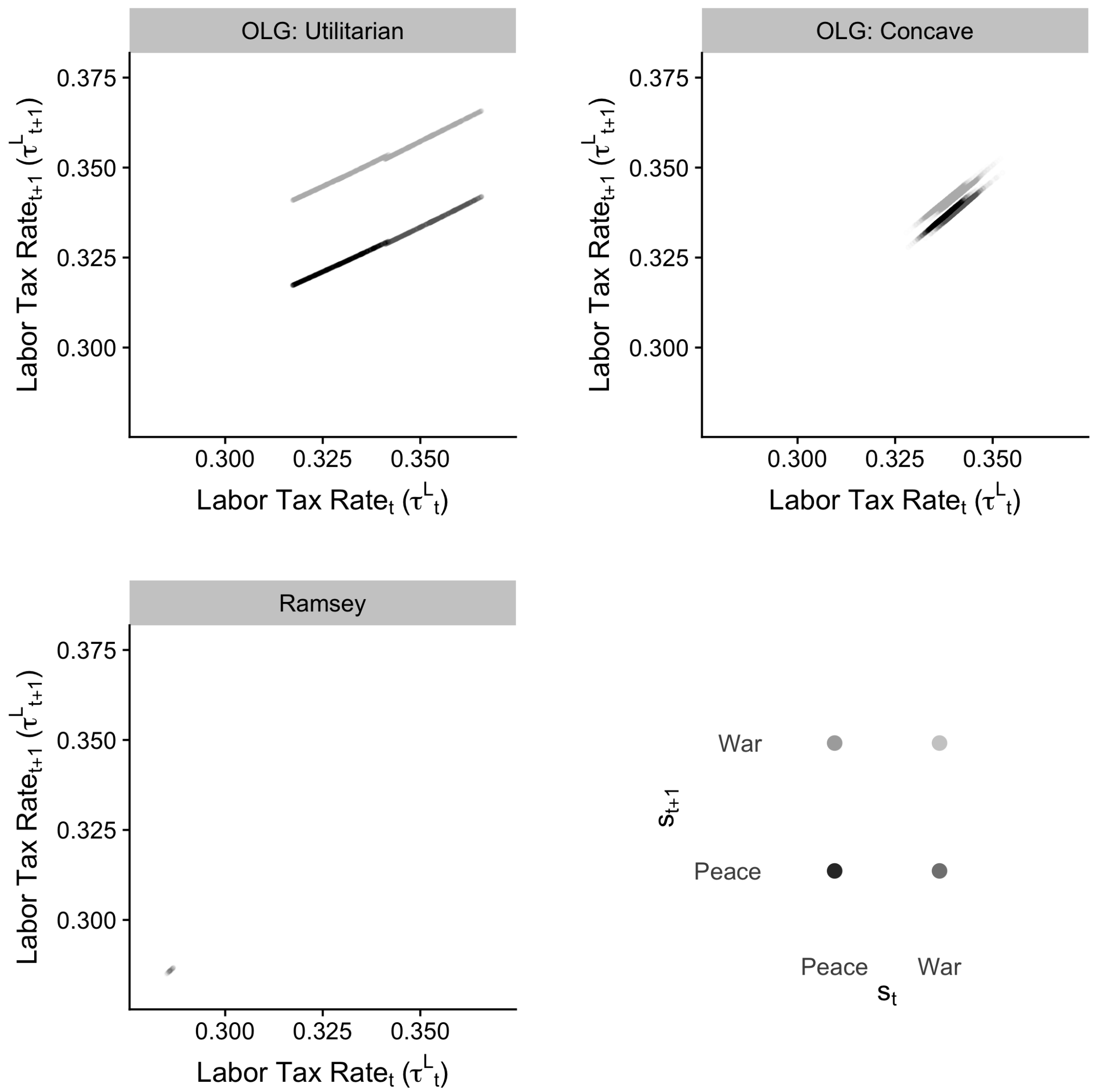


Figure A.4: This figure presents the results of a random simulation of the complete markets model of stochastic government expenditure when utility is isoelastic with $\sigma = 1/4$. The model is simulated over 10000 periods (after a 100-period startup period that is excluded), and each point represents one period. The concave planner has $\zeta = -8$. The labor tax rate at date t is on the horizontal axis, and the labor tax rate at date $t+1$ is on the vertical axis. The different shades of gray represent different sequences of shocks received at date t (s_t) and date $t+1$ (s_{t+1}). The Ramsey model was started with debt equal to -3.5% of nonstochastic steady state GDP, which corresponds to the level of government debt in the OLG model after 100 periods of low government spending; initial conditions are irrelevant for the OLG models.

utility or inequality in consumption equivalents. I will not attempt that exercise here.

Incomplete Markets. The incomplete markets restriction might seem to be irrelevant in this case, as there are two assets (capital and risk free debt) and two states, while debt as a percentage of GDP was sufficiently small in the complete markets case so as to not violate any imposed debt limits. However, recall that if utility is log-log, then there is constant labor supply if there are no lump-sum transfers and, absent productivity shocks, this implies that capital is actually riskless; thus the available assets do *not* span the state space in the log-log case. Additionally, while households may have access to complete markets over shocks received in old age, the government does not; the government may only directly interact with the risk free debt market, and faces position limits in even that market. The government may tax capital. However, taxing even the gross return to capital is not equivalent to pure ownership of the capital since there is no expensing of investment.

Figures A.5 and A.6 respectively show the behavior of several variables across state transitions, and the autocorrelation structure of labor taxes, for the log-log case. Perhaps most striking is that the graphs are qualitatively quite similar to the equivalent complete markets graphs: The transitions exhibit most of the same properties, and the scatterplot exhibits the same autocorrelation structure within combinations of spending and lagged spending, while having these clusters of points spread out across the graph. This similarity might initially seem surprising, but introducing the OLG structure to the model in a sense introduces a substantial amount of market incompleteness, even if the model features recursively complete markets. This is because, while the government *can* insure itself completely against government spending shocks, it *chooses* not to due to both the efficiency and equity concerns discussed in the main text. Put another way, individuals cannot insure themselves against the state of the world into which they are

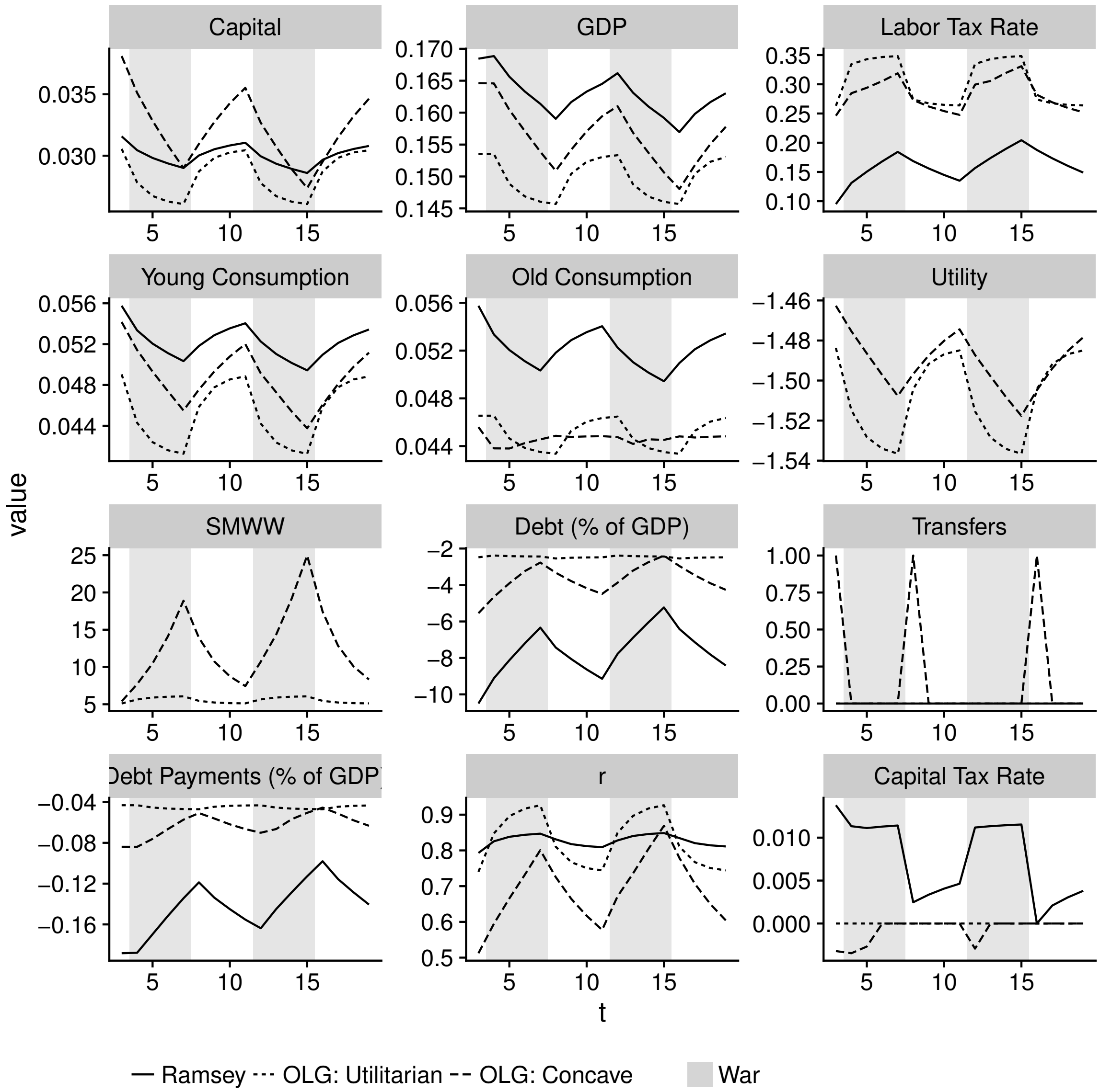


Figure A.5: This figure captures the salient features of state transitions in the incomplete markets model of stochastic government expenditure when utility is log-log. Shading marks periods of high government spending ("wars"). The concave planner has $\zeta = 24$. The model was first allowed to run for 100 periods in state 1 to arrive at a steady state. The Ramsey model was started with debt equal to -2.4% of nonstochastic steady state GDP, which corresponds to the level of government debt in the OLG model after 100 periods of low government spending; initial conditions are irrelevant for the OLG models.

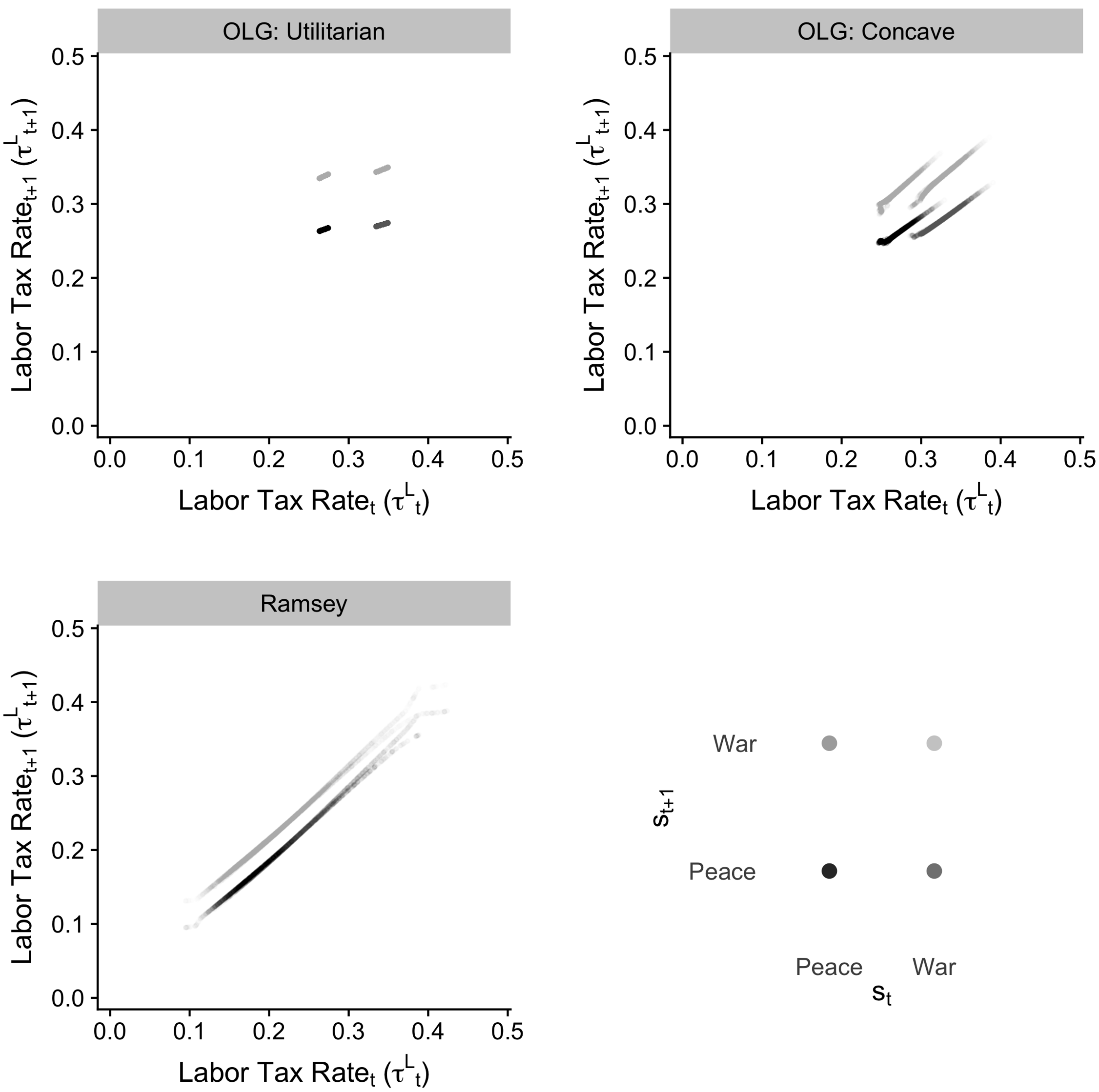


Figure A.6: This figure presents the results of a random simulation of the incomplete markets model of stochastic government expenditure when utility is log-log. The model is simulated over 10000 periods (after a 100-period startup period that is excluded), and each point represents one period. The concave planner has $\zeta = 24$. The lagged labor tax rate is on the horizontal axis, and the current labor tax rate is on the vertical axis. The different shades of gray represent different sequences of shocks received at date t (s_t) and date $t+1$ (s_{t+1}).

born, and the government cannot insure itself against shocks using the entire collection of agents as counterparties, which makes markets substantively incomplete.

The one qualitative difference from complete markets is that the government makes use of some lump-sum transfers in old age. Specifically, upon exiting a war, the planner compensates the old with a very small lump-sum transfer. The reason that transfer is not larger is that the efficiency cost of these transfers is quite large; they reduce labor supply, through the income effect, in the previous period. A much more nuanced discussion of old age transfers and state contingent debt can be found in the sections on productivity shocks and trend productivity growth—applications that better lend themselves to study of this phenomenon.

Figures A.7 and A.8 respectively show the behavior of several variables across state transitions, and the autocorrelation structure of labor taxes, if the IES is 1/4 and markets are incomplete. The concave planner has $\zeta = -8$. Likewise, Figures A.9 and A.10 respectively show the behavior of several variables across state transitions, and the autocorrelation structure of labor taxes, if the IES is 4 and markets are complete. The concave planner has $\zeta = 8$. Again, these are quite qualitatively similar to their log-log equivalents, with just a couple of changes. Lump-sum transfers to the old do play a somewhat larger role, even in percentage terms and even for the utilitarian planner—especially for low IES. This stems from the issue mentioned above: Utilitarian (or others) planners' preference for equality in dollar terms is not constant, but depends on households' risk aversion. Since these higher σ households have higher risk aversion, even a utilitarian planner will have a stronger motive for redistribution and therefore compel the old to hold a larger share of risk. The other alteration is that, for the first time, we see that capital is non-negligibly taxed (or subsidized) by the concave planner. This occurs with the opposite sign that would be predicted by the hedging term, and therefore suggests that income effects dominate the capital tax discussion.

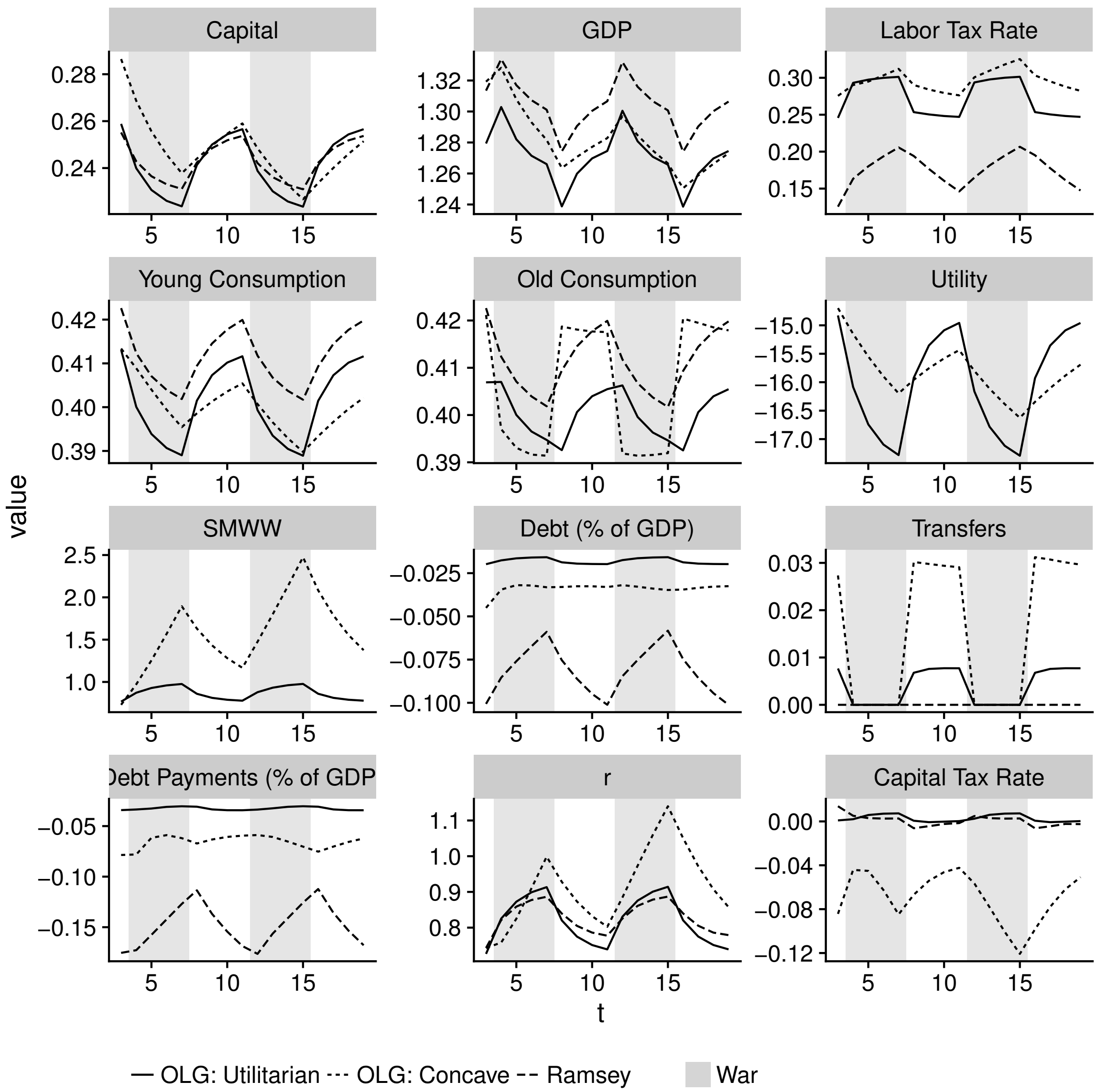


Figure A.7: This figure captures the salient features of state transitions in the incomplete markets model of stochastic government expenditure when utility is isoelastic with $\sigma = 4$. Shading marks periods of high government spending ("wars"). The concave planner has $\zeta = -8$. The model was first allowed to run for 100 periods in state 1 to arrive at a steady state.

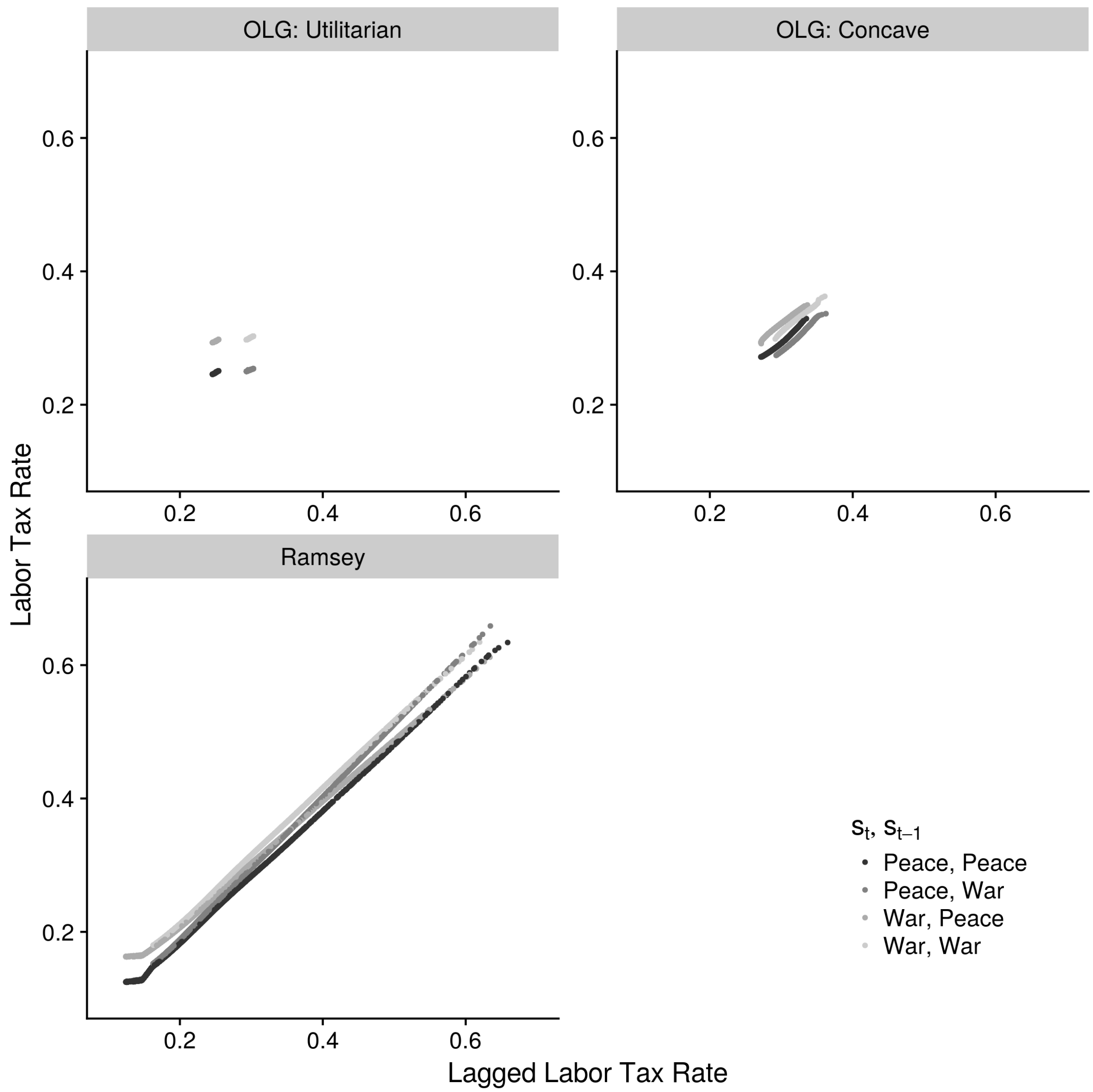


Figure A.8: This figure presents the results of a random simulation of the incomplete markets model of stochastic government expenditure when utility is isoelastic with $\sigma = 4$. The model is simulated over 10000 periods (after a 100-period startup period that is excluded), and each point represents one period. The concave planner has $\zeta = -8$. The lagged labor tax rate is on the horizontal axis, and the current labor tax rate is on the vertical axis. The different shades of gray represent different sequences of shocks received at date t (s_t) and date $t + 1$ (s_{t+1}).

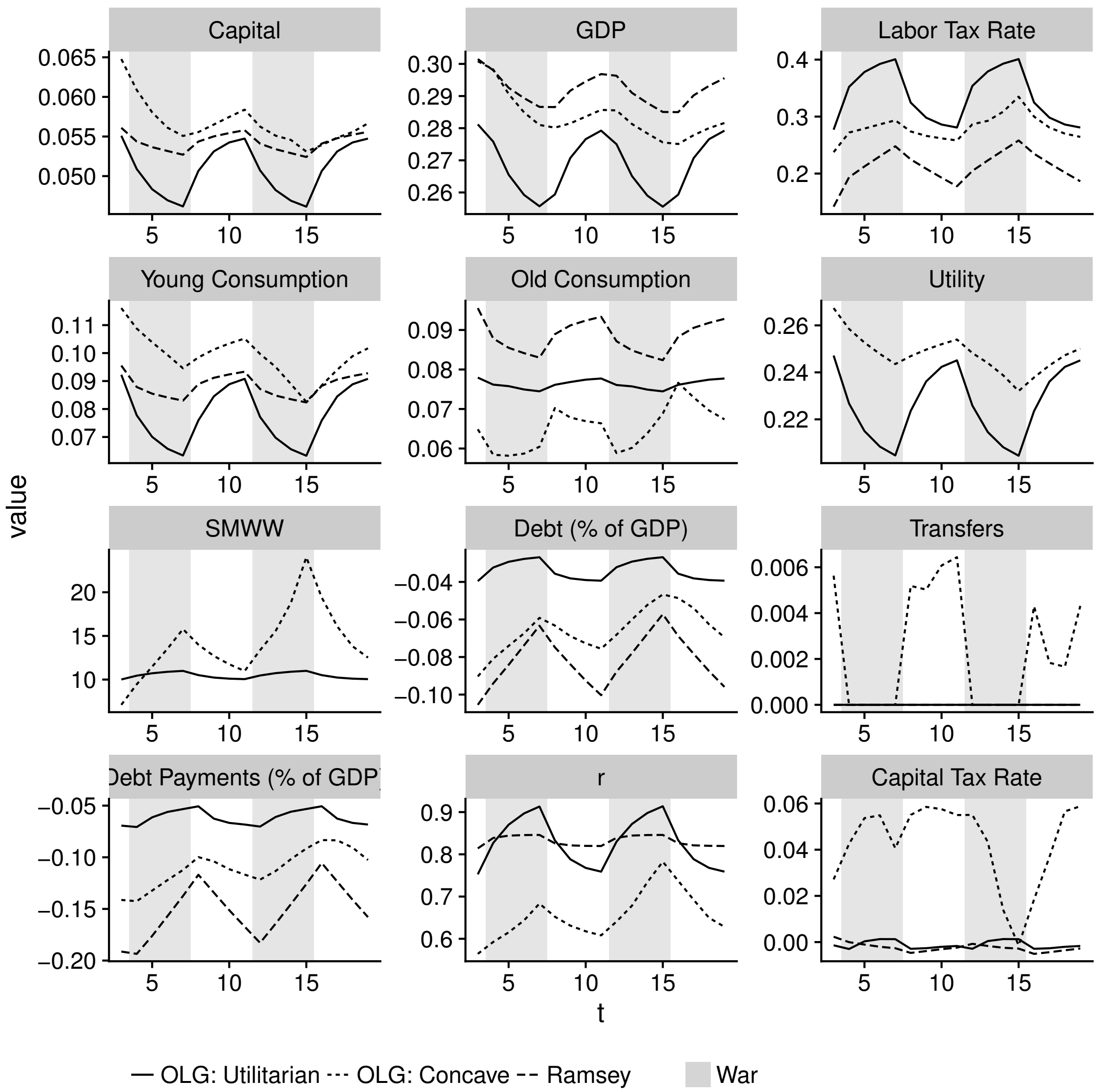


Figure A.9: This figure captures the salient features of state transitions in the incomplete markets model of stochastic government expenditure when utility is isoelastic with $\sigma = 1/4$. Shading marks periods of high government spending ("wars"). The concave planner has $\zeta = 8$. The model was first allowed to run for 100 periods in state 1 to arrive at a steady state.

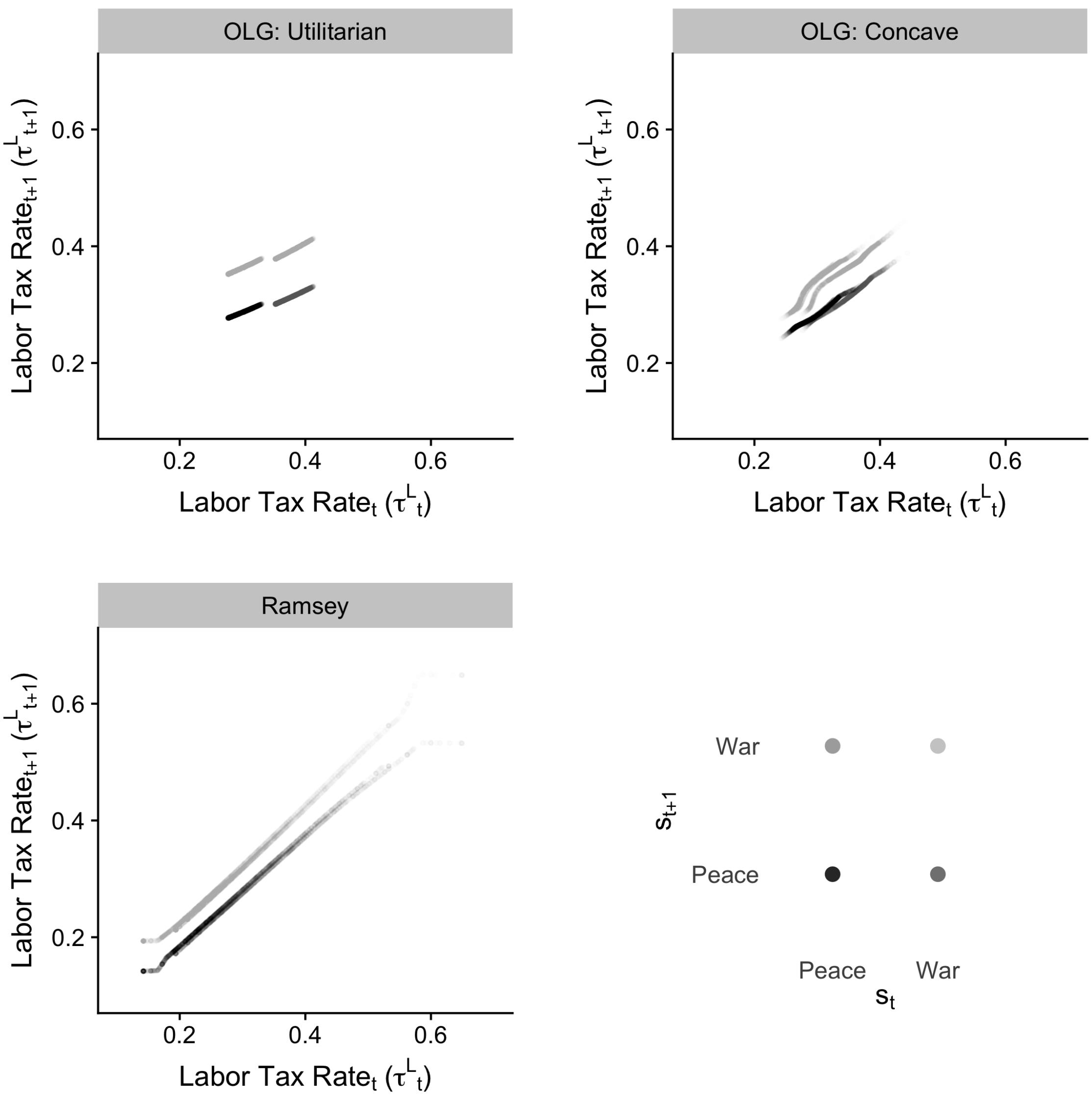


Figure A.10: This figure presents the results of a random simulation of the incomplete markets model of stochastic government expenditure when utility is isoelastic with $\sigma = 1/4$. The model is simulated over 10000 periods (after a 100-period startup period that is excluded), and each point represents one period. The concave planner has $\zeta = -8$. The lagged labor tax rate is on the horizontal axis, and the current labor tax rate is on the vertical axis. The different shades of gray represent different sequences of shocks received at date t (s_t) and date $t + 1$ (s_{t+1}).

A.5.2 Productivity Shocks

I limit discussion to complete markets models. As mentioned in the text, this seems the more natural assumption when discussing productivity shocks.[7] Figure A.11 shows the behavior of several variables across state transitions if the IES is 1/4 and markets are complete. The concave planner has $\zeta = -2$. Likewise, Figure A.12 shows the same when the IES is 4. The concave planner has $\zeta = 8$. The primary difference is the fact that income effects are so strong when $\sigma = 4$ that taxes are actually higher on less well-off generations. Additionally, defined constribution SS rises as a percentage of GDP for the utilitarian planner when the economy enters a recession, as helping retirees with their very high risk aversion takes priority over insuring the unborn. Policies are more variable when $\sigma = 1/4$, since the economy is closer to Ricardian and thus redistributive policy is more palatable.

A.6 Definitions and Proofs from Section 1.6

Definition A.6.1 (Policy Equilibrium—More than Two Generations) *Given a set of after-tax assets held at the beginning of time, $\{a^j(s_0)\}_{j=2}^{J_2}$, a policy equilibrium is a collection of*

- *policies* $\{T_0^j\}_{j\geq 2}, \{T(s^t), \tau^L(s^t), \tau^K(s^{t+1})\}_{t\geq 0}$
- *prices* $\{q(s^{t+1}), w(s^t), R^K(s^t)\}_{t\geq 0}$
- *an allocation* $\{\{c^j(s^t)\}_{j=1}^{J_2}, \{\ell^j(s^t)\}_{j=1}^{J_2}, \{a^{j+1}(s^{t+1})\}_{j=1}^{J2-1}, k(s^t)\}_{t\geq 0}$
- *and government debt* $\{b(s^{t+1})\}_{t\geq 0}$

[7] Additionally, the caveats mentioned in the first footnote to this Appendix apply especially strongly here, making incomplete markets results highly dubious.

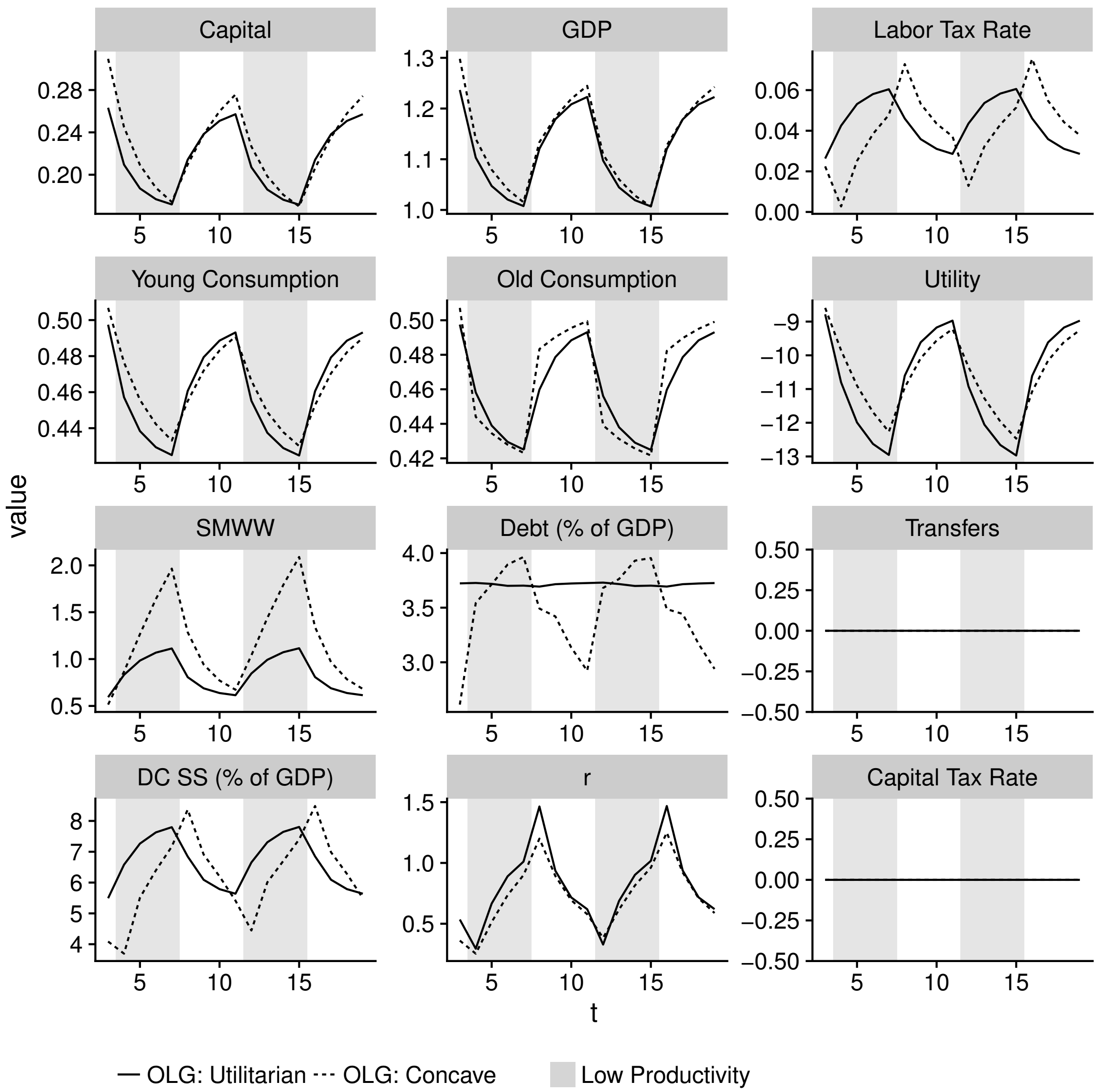


Figure A.11: This figure captures the salient features of state transitions in the complete markets model of stochastic productivity when utility is isoelastic with $\sigma = 4$. Shading marks periods of low productivity. The concave planner has $\zeta = -2$. The model was first allowed to run for 100 periods in state 1 to arrive at a steady state.

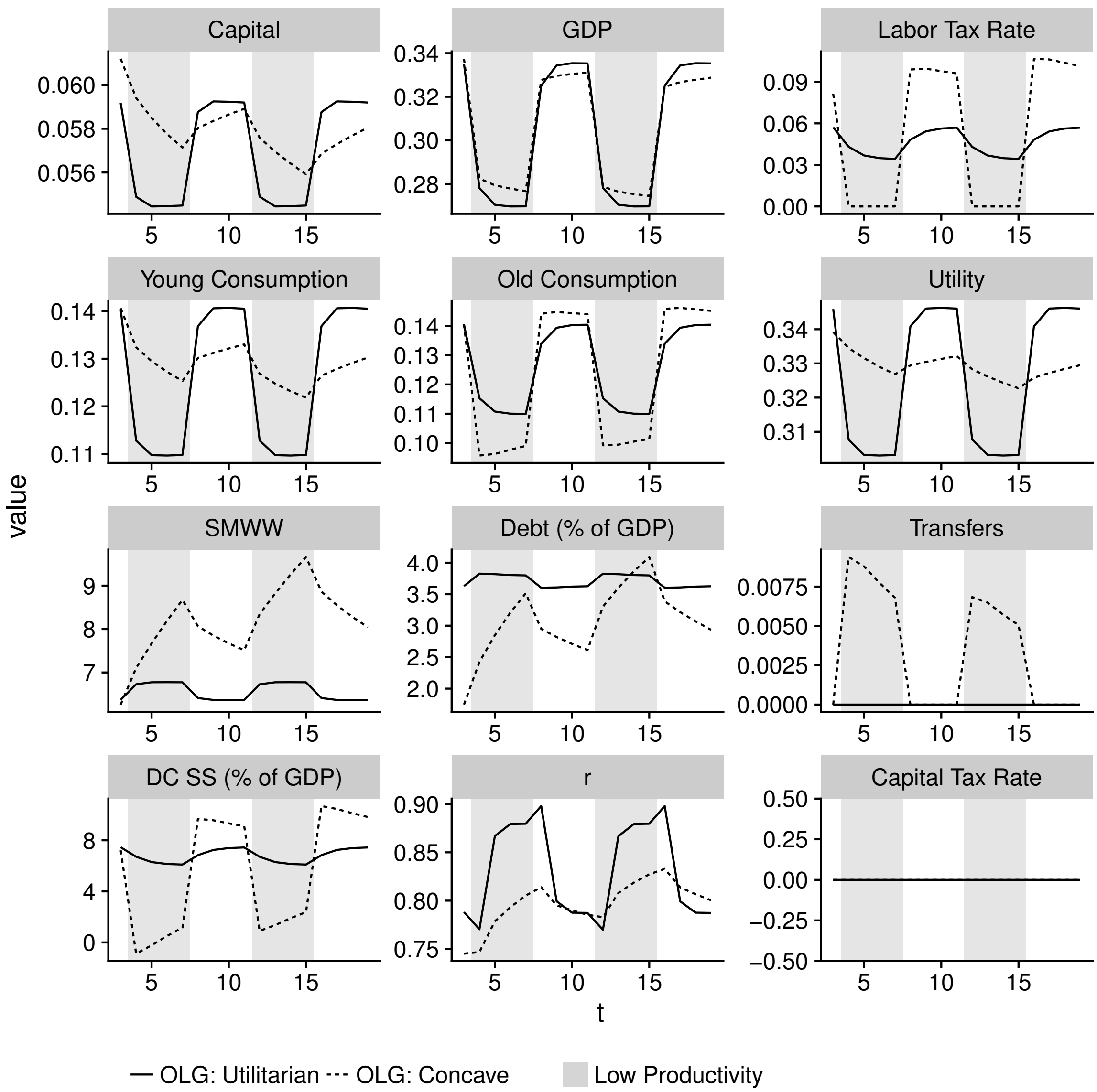


Figure A.12: This figure captures the salient features of state transitions in the complete markets model of stochastic productivity when utility is isoelastic with $\sigma = 1/4$. Shading marks periods of low productivity. The concave planner has $\zeta = 8$. The model was first allowed to run for 100 periods in state 1 to arrive at a steady state.

such that at all histories s^t, $t \geq 0$

- *The resource constraint is satisfied at all* $t \geq 0$*:*

$$\sum_{j=1}^{J_2} c^j(s^t) + g(s_t) + k(s^t) \leq F\left(k(s^{t-1}), \sum_{j=1}^{J_1} \ell(s^t)\right) \tag{A.2}$$

- *The government's budget constraint is satisfied at all* $t \geq 0$*:*

$$\begin{aligned} b(s^t) + g(s_t) + T(s^t) \leq & \sum_{s_{t+1}} q(s^{t+1})b(s^{t+1}) + \tau^L(s^t)w(s^t)\sum_{j=1}^{J_1}\ell^j(s^t) \\ & + \tau^K(s^t)R^K(s^t)k(s^{t-1}) \end{aligned} \tag{A.3}$$

- *Households' budget constraints are satisfied:*

$$c^j(s_0) + \sum_{s_1} q(s^1)a^{j+1}(s^1) \leq w(s_0)\ell^j(s_0) + a^j(s_0) + T_0^j \quad j = 2, ...J_1$$

$$c^j(s_0) + \sum_{s_{t+1}} q(s^1)a^{j+1}(s^1) \leq a^j(s_0) + T_0^j \quad j = J_1 + 1, ...J_2 - 1$$

$$c^{J2}(s_0) \leq a^{J2}(s_0) + T_0^{J2}$$

$$c^1(s^t) + \sum_{s_{t+1}} q(s^{t+1})a^2(s^{t+1}) \leq w(s^t)\ell^1(s^t) + T(s^t)$$

$$c^j(s^t) + \sum_{s_{t+1}} q(s^{t+1})a^{j+1}(s^{t+1}) \leq w(s^t)\ell^j(s^t) + a^j(s^t)\, j = 2, ...J_1$$

$$c^j(s^t) + \sum_{s_{t+1}} q(s^{t+1})a^{j+1}(s^{t+1}) \leq a^j(s^t)\, j = J_1 + 1, ...J_2 - 1$$

$$c^{J2}(s^t) \leq a^{J2}(s^t)$$

- *Households optimize subject to their budget constraint, taking prices and policies as given*

- *Firms optimize, taking prices as given*

- *The state-contingent asset markets clear:*

$$\sum_{j=2}^{J_2} a^j(s^{t+1}) = b(s^{t+1}) + (1-\tau^K(s^{t+1}))R^K(s^{t+1})k(s^t)$$

- *The markets for capital and labor clear*

- *The no-arbitrage condition between capital and state-contingent assets holds:*

$$\sum_{s_{t+1}} q(s^{t+1})R^K(s^{t+1})(1-\tau^K(s^{t+1})) = 1$$

Definition A.6.2 (Implementable Allocation, >2 Generations) *An allocation is implementable if there exists a policy equilibrium of which it is apart. More importantly a welfare-relevant allocation—an allocation, ignoring $a^j(s^t)$— is implementable if there exists a policy equilibrium of which it is a part.*

Proposition A.6.1 (Implementability, >2 Generations) *Given a set of after-tax assets held at the beginning of time, $\{a^j(s_0)\}_{j=2}^{J_2}$, a set*

$$\{\{c^j(s^t)\}_{j=1}^{J_2}, \{\ell^j(s^t)\}_{j=1}^{J_2}, k(s^t)\}_{t\geq 0}$$

is part of an implementable allocation if and only if

- *the resource constraint is satisfied*

- *the implementability condition is satisfied for all generations not already alive*

$$\mathbb{E}_t \sum_{j=1}^{J_2} \beta^{j-1} u_c^j(s^{t+j-1})c^j(s^{t+j-1}) + \mathbb{E}_t \sum_{j=1}^{J_1} \beta^{j-1} u_\ell^j(s^{t+j-1})\ell^j(s^{t+j-1}) \geq 0 \quad \forall t \geq 0 \tag{A.4}$$

- *the implementability condition is satisfied for all generations already alive*

$$\mathbb{E}_t \sum_{j'=j}^{J_2} \beta^{j'-j} u_c^{j'}(s^{j'-j}) c^{j'}(s^{j'-j}) + \mathbb{E}_t \sum_{j'=j}^{J_1} \beta^{j'-j} u_\ell^{j'}(s^{j'-j}) \ell^{j'}(s^{j'-j}) \geq u_c^j(s_0) a^j(s_0)$$

$$\forall j = 2, ..., J_2 \tag{A.5}$$

- *there exists a set of market weights* $\{\varphi(s^t)\}_{t=-J_2+1}^{\infty}$ *such that, after defining* $C(s^t) \equiv \sum_{j=1}^{J_2} c^j(s^t)$ *and* $L(s^t) \equiv \sum_{j=1}^{J_1} \ell^j(s^t)$*:*

$$\{c^j(s^t)\}_{j=1}^{J_2}, \{\ell^j(s^t)\}_{j=1}^{J_1} \in argmax \sum_{j=1}^{J_1} \varphi(s^{t-j+1}) u(c_j, \ell_j) + \sum_{j=J_1+1}^{J_2} \varphi(s^{t-j+1}) u(c_j, 0)$$

$$s.t. \sum_{j=1}^{J_2} c_j \leq C(s^t)$$

$$\sum_{j=1}^{J_1} \ell_j \geq L(s^t)$$

Proof. I will prove the "if" direction first, by constructing a policy equilibrium. First define prices as follows:

$$w(s^t) = F_L\left(k(s^{t-1}), \sum_{j=1}^{J_1} \ell^j(s^t)\right)$$

$$R^K(s^t) = F_K\left(k(s^{t-1}), \sum_{j=1}^{J_1} \ell^j(s^t)\right)$$

$$q(s^{t+1}) u_c^1(s^t) = \beta Pr(s^{t+1}|s^t) u_c^2(s^{t+1})$$

Then define policies as follows:

$$w(s^t)(1 - \tau^L(s^t)) = -\frac{u_\ell^1(s^t)}{u_c^1(s^t)}$$

$$u_c^1(s^t) = \beta(1 - \kappa(s^t)) \mathbb{E}_t[u_c^2(s^{t+1}) R^K(s^{t+1}]$$

$$\tau^K(s^{t+1}) = \kappa(s^t) \quad \forall s^{t+1} \succ s^t$$

Most importantly, I must show that the market weights ensure that that other generations' (i.e., not the youngest's) first order conditions are satisfied with the proposed allocation and these prices. Define $\psi(s^t)$ and $\lambda(s^t)$ as the multipliers on the consumption and labor constraints from the intra-period allocation problem. Then for all $j = 1, ..., J_1$, we have

$$-\frac{u_l^j(s^t)}{u_c^j(s^t)} = \frac{\lambda(s^t)/\varphi(s^{t-j+1})}{\psi(s^t)/\varphi(s^{t-j+1})} = \frac{\lambda(s^t)}{\psi(s^t)}$$

which shows that all households' labor-leisure first order condition is satisfied at the same net-of-tax wage. Since we chose the net-of-tax wage to satisfy the youngest household's labor-leisure first order condition, all other households must be optimizing as well. Similarly, the Euler equation with respect to a state-contingent asset paying off at s^{t+1} for a household of age j is

$$\begin{aligned}
q(s^{t+1})u_c^j(s^t) &= \beta Pr(s^{t+1}|s^t)u_c^{j+1}(s^{t+1}) \\
q(s^{t+1}) &= \frac{\beta Pr(s^{t+1}|s^t)u_c^{j+1}(s^{t+1})}{u_c^j(s^t)} \\
&= \frac{\beta Pr(s^{t+1}|s^t)\psi(s^{t+1})/\varphi(s^{t-j+1})}{\psi(s^t)/\varphi(s^{t-j+1})} \\
&= \beta Pr(s^{t+1})\frac{\psi(s^{t+1})}{\psi(s^t)}
\end{aligned}$$

which again shows that all households' Euler equations are satisfied at the same asset price, and since we chose $q(s^{t+1})$ such that the youngest household's Euler equation is satisfied, so must all other households'. For households to be optimizing given their budget constraint, all that remains is to show that their budget constraints are satisfied. But this is enforced by the implementability conditions for an appropriately chosen $T(s^t)$.

Finally, we must show that asset markets clear, the no arbitrage condition with respect to capital is satisfied, and the government's budget constraint is satisfied. Given the capital tax chosen above, the no arbitrage condition follows immediately. Asset holdings can be defined as a forward looking variable. Specifically, set

$$u_c^j(s^t)a^j(s^t) = \mathbb{E}_t \sum_{j'=j}^{J_2} \beta^{j'-j} u_c^{j'}(s^{t+j'-j})c^{j'}(s^{t+j'-j}) + \mathbb{E}_t \sum_{j'=j}^{J_1} \beta^{j'-j} u_\ell^{j'}(s^{t+j'-j})\ell^{j'}(s^{t+j'-j}).$$

Substituting this definition into the right hand side leaves

$$\begin{aligned} u_c^j(s^t)a^j(s^t) =& u_c^j(s^t)c^j(s^t) + \mathbb{1}_{j\leq J_1} u_\ell^j(s^t)\ell^j(s^t) \\ &+ \mathbb{1}_{j<J_2}\beta \sum_{s_{t+1}} Pr(s^{t+1}|s^t)u_c^{j+1}(s^{t+1})a^{j+1}(s^{t+1}) \end{aligned}$$

Dividing through by $u_c^j(s^t)$ yields

$$a^j(s^t) = c^j(s^t) - \mathbb{1}_{j\leq J_1} w(s^t)(1-\tau^L(s^t))\ell^j(s^t) + \mathbb{1}_{j<J_2} \sum_{s_{t+1}} q(s^{t+1})a^{j+1}(s^{t+1}) \qquad \text{(A.6)}$$

The equivalent expression for the youngest generation is

$$T(s^t) = c^1(s^t) - w(s^t)(1-\tau^L(s^t))\ell^1(s^t) + \sum_{s_{t+1}} q(s^{t+1})a^2(s^{t+1}) \qquad \text{(A.7)}$$

Defining $b(s^t) = \sum_{j=2}^{J_2} a^j(s^t) - (1-\tau^K(s^t))R^K(s^t)k(s^{t-1})$, which ensures the that the asset markets clear, and summing these last two equations across all j yields

$$\left\{\begin{array}{c} T(s^t)+b(s^t) \\ +(1-\tau^K(s^t))R^K(s^t)k(s^{t-1}) \end{array}\right\} = \left\{\begin{array}{c} \sum_{j=1}^{J_2} c^j(s^t) \\ -(1-\tau^L(s^t))w(s^t)\sum_{j=1}^{J_1}\ell^j(s^t) \\ +\sum_{s_{t+1}} q(s^{t+1})(b(s^{t+1}) \\ +(1-\tau^K(s^{t+1})R^K(s^{t+1})k(s^t)) \end{array}\right\}$$

$$\left\{\begin{array}{c} T(s^t)+b(s^t) \\ +F\left(k(s^{t-1}),\sum_{j=1}^{J_1}\ell^j(s^t)\right) \end{array}\right\} = \left\{\begin{array}{c} \tau^K(s^t)R^K(s^t)k(s^{t-1})+ \\ \tau^L(s^t)w(s^t)\sum_{j=1}^{J_1}\ell^j(s^t) \\ +\sum_{j=1}^{J_2} c^j(s^t)+k(s^t) \\ +\sum_{s_{t+1}} q(s^{t+1})b(s^{t+1}) \end{array}\right\}$$

$$\begin{aligned} T(s^t)+b(s^t)+g(s_t) =& \tau^K(s^t)R^K(s^t)k(s^{t-1})+\tau^l(s^t)w(s^t)\sum_{j=1}^{J_1}\ell^j(s^t) \\ &+\sum_{j=1}^{J_2} c^j(s^t)+k(s^t)+\sum_{s_{t+1}} q(s^{t+1})b(s^{t+1}) \\ &+g(s_t)-F\left(k(s^{t-1}),\sum_{j=1}^{J_1}\ell^j(s^t)\right) \\ \leq& \tau^K(s^t)R^K(s^t)k(s^{t-1})+\tau^L(s^t)w(s^t)\sum_{j=1}^{J_1}\ell^j(s^t) \\ &+\sum_{s_{t+1}} q(s^{t+1})b(s^{t+1}) \end{aligned}$$

Thus, the “if” direction has been proven.

Proving the “only if” direction requires showing that every policy equilibrium satisfies the conditions. The resource constraint is direct, and the implementability conditions follow directly from the household’s budgets and first order conditions. I must show that there exists a set of market weights with the properties specified. I do so by construction.

First define $\varphi(-J_2+1)=1$. For any generation already alive, set $\varphi(-j+1)$

such that $\frac{u_c^j(s_0)}{u_c^{J_2}(s_0)} = \frac{\varphi(-J_2+1)}{\varphi(-j+1)}$. For any generation not already alive, set $\varphi(s^t)$ such that $\frac{u^1(s^t)}{u^2(s^t)} = \frac{\varphi(s^{t-1})}{\varphi(s^t)}$. This, by construction, enforces the first order conditions of the intra-period allocation problem with respect to consumption,

$$\varphi(s^{t-j+1})u_c^j(s^t) = \varphi(s^{t-j'+1})u_c^{j'}(s^t).$$

Multiplying both sides by $(1-\tau^L(s^t))w(s^t)$ for all pairs of young generations yields

$$-\varphi(s^{t-j+1})u_\ell^j(s^t) = -\varphi(s^{t-j'+1})u_\ell^{j'}(s^t)$$

which is the first order condition with respect to labor of the intra-period allocation problem. Thus, with $C(s^t)$ and $L(s^t)$ suitably defined, the intra-period allocation problem indeed yields the actual allocation for these φs, and the "only if" direction is proven. ■

The planner's problem thus consists of choosing only aggregates, plus the Pareto weights. Any given period, he may only choose a single Pareto weight—that of the youngest generation; all other Pareto weights are simply moved up one "slot." The resulting recursive problem for a utilitarian planner is as follows, properly interpreting c^j and ℓ^j as functions of $C, L, \{\varphi_j\}_{j=1}^{J_2}$:

Problem A.6.1 (Planner's Problem, > 2 Generations) *Define*

$$V(k, \{\theta_j\}_{j=1}^{J_2-1}, \{\varphi_j\}_{j=1}^{J_2-1}) =$$

$$\max_{C, L, k', \varphi_1', \{\theta_{s'}'^j\}_{j=1,s'}^{J_2-1}} \quad \sum_{j=1}^{J_1} u(c_j, \ell_j) + \sum_{j=J_1+1}^{J_2} u(c_j, 0)$$

$$+ \beta \mathbb{E} V(k', \{\theta_{s'}'^j\}_{j=1}^{J_2-1}, (\varphi_1', \varphi_1, ..., \varphi_{J2-2}))$$

$$\text{s.t.}$$

$$C + g(s_-) + k' \leq F(k, L)$$

$$u_c(c_1, \ell_1)c_1 + u_\ell(c_1, \ell_1)\ell_1 + \beta \mathbb{E}[\theta_{s'}'^1 | s] \geq 0$$

$$u_c(c_j, \ell_j)c_j + u_\ell(c_j, \ell_j)\ell_j + \beta \mathbb{E}[\theta_{s'}'^j | s] \geq \theta^{j-1} \quad j = 2, ..., J_1$$

$$u_c(c_j, 0)c_j + \beta \mathbb{E}[\theta_{s'}'^j | s] \geq \theta^{j-1} \quad j = J_1 + 1, ..., J_2 - 1$$

$$u_c(c_{J_2}, 0)c_{J_2} \geq \theta^{J_2-1}$$

Then the planner solves

$$\max_{C, L, k', \{\varphi_j\}_{j=1}^{J_2}, \{\theta_{s'}'^j\}_{j=1,s'}^{J_2-1}} \quad \sum_{j=1}^{J_1} u(c_j, \ell_j) + \sum_{j=J_1+1}^{J_2} u(c_j, 0)$$

$$+ \beta \mathbb{E} V(k', \{\theta_{s'}'^j\}_{j=1}^{J_2-1}, (\varphi_1, \varphi_2, ..., \varphi_{J2-1}))$$

$$\text{s.t.}$$

$$C + g(s_-) + k' \leq F(k, L)$$

$$u_c(c_1, \ell_1)c_1 + u_\ell(c_1, \ell_1)\ell_1 + \beta \mathbb{E}[\theta_{s'}'^1 | s] \geq 0$$

$$u_c(c_j, \ell_j)c_j + u_\ell(c_j, \ell_j)\ell_j + \beta \mathbb{E}[\theta_{s'}'^j | s] \geq u_c(c_j, \ell_j)a_j \quad j = 2, ..., J_1$$

$$u_c(c_j, 0)c_j + \beta \mathbb{E}[\theta_{s'}'^j | s] \geq u_c(c_j, 0)a_j \quad j = J_1 + 1, ..., J_2 - 1$$

$$c_{J_2} \geq a_{J_2}$$

Since the magnitude of $\{\varphi_j\}_{j=1}^{J_2}$ is irrelevant, the planner can always normalize $\phi_1 = 1$ before moving on to the next period, thus keeping the state space contained. An incomplete markets version, and/or a non-utilitarian version, can be defined similarly. However, both require substantially more state variables than the complete markets, utilitarian version, and so I will not simulate them.

Appendix B

Appendix to Chapter 2

B.1 Data: Technical Details

Details of DataQuick Cleaning

As discussed in the text, I mostly follow Guren (2018) for cleaning the DataQuick data. This primarily involves dropping transactions that are not sales (such as refinances), duplicate transactions that cannot be easily consolidated, and "partial consideration sales, partial sales, group sales, and splits." I make one major change. First, I am much more conservative regarding dropping of transactions that appear not to be arm's length. Guren (2018) flagged all transactions that (a) are flagged by DataQuick as not arm's length, (b) involve too much similarity between buyer and seller names, or (c) have a price of zero dollars. Clerical review, however, shows that the DataQuick flag is likely to have been miscoded in a nontrivial number of jurisdictions. In most jurisdictions, a relatively small percentage of transactions are flagged as not arm's length, and those observations tend to involve no simultaneous mortgages. However, in some jurisdictions (for example, all of Texas), those attributes are reversed, which suggests reverse coding; furthermore, in many other jurisdictions, the attributes are relatively similar between the transactions that are flagged and not flagged, suggesting perhaps that the transactions were coded correctly in some cases, and reverse coded in others. In my opinion, the most conservative approach is to include observations that were flagged by DataQuick as arm's length transfers; fortunately, matching to mortgage data will eliminate those transactions that are unaccompanied by a mortgage, correcting a substantial fraction of the problem. Observations that Guren (2018) flagged as likely not arm's length, but which DataQuick did not, I continue to drop; these involve similar buyer and seller names, or zero price.

Details of Home Mortgage Disclosure Act Data

HMDA LAR data can be obtained from the National Archives (identifier 2456161) from 1981–2013. I focus on the period 1992–2012. The data is distributed in fixed width format, one file per year, and I used the accompanying formatting records to split the data into columns. I then split the data into counties, but aggregated across years for each county. The names of the lending institutions are then merged in from similarly downloaded and split Transmittal Sheet (TS) files. The format of the data changed in 1998 and again in 2004, so this needed to be undertaken in three parts. After selecting relevant columns, I gave each application a unique identifier within the county level dataset.

Census Demographic Data

Census data on racial composition and income at the block group level was downloaded from NHGIS (Manson et al., 2017). In particular, to establish the *levels* of these variables, I used Tables NP004A (Population by Hispanic or Latino and Not Hispanic or Latino), NP007A (Population by Race) from Summary File 1b (100% Data), as well as Table NP053A (Median Household Income in 1999) from Summary File 3a (Sample-Based Data), all from the 2000 Decennial Census, as it represents roughly the midpoint of my analysis. To establish *changes* in the racial composition of neighborhoods, I used Time Series Tables CM1 (Persons by Race) and CP4 (Persons by Hispanic or Latino Origin). These data are standardized to 2010 Census geographies by using block level data from the original year and aggregating according to 2010 boundaries; more details can be found in the NHGIS documentation. Unfortunately, NHGIS cannot standardize the income variables, since those are not colleted at the block level. After collecting these data, the demographic variables I use are created in the obvious way.

Equality of Opportunity Project Data

Data on the demographics of commuting zones, including racial composition, segregation, income, income inequality, and others, are drawn from the Equality of Opportunity Project, particularly from Chetty et al. (2014) and Chetty et al. (2018). These also map 1990 counties into commuting zones.

Geographical Data

Data on Census geography was downloaded using the "tigris" R package, which in turn packages data from Census TIGER/Line®Shapefiles. Each geography is represented as a "Simple Features Data Frame" using the "sf" R package—a package that provides R with many traditional GIS capabilities. This data structure includes both data (in this case, the identifiers for the geographies) and the set of geometries that comprise the geography.

Merging

Merging these data required several steps. Each step of the merge was performed one county at a time as recorded by DQ, with the results aggregated at the very end. Unfortunately, county boundaries are not perfectly uniform across geography vintages, and so a few adjustments were necessary using comments from the Census Bureau website.[1] Care, therefore, had to be taken when matching to HMDA LAR or Census data, which was gathered one county at a time as recorded by *that* dataset, which is often from 1990 or 2010. Fortunately, except for the Dade County renaming, none of these changes affected very many properties.

The broad merging process can be summarized as follows:

1. Merge DQ history data to DQ assessor data using the unique DQ identifier

2. Merge DQ history data to Census geographies of various vintages using latitude and longitude of home, Census TIGER/Line®Shapefiles, and the R "sf" package's GIS capabilities

3. Merge this geo-allocated data to HMDA LAR data by requiring perfect matching on Census tract (of the appropriate vintage), loan amount, and first letter of lending instiution, and a probabilistic match on the entire name of the lending institution using the R "RecordLinkage" package

4. Merge the geo-allocated data to Census demographic data using the Census geography identifiers

[1] For example, Broomfield County, CO, was created during the 2000s from regions of several neighboring counties, though its FIPS code does appear in the DQ data that otherwise uses 2000 Census geographies. Additionally, several counties were absorbed during the 1990s and 2000s, and Dade County, FL, changed its name (and therefore FIPS code) during the 1990s to Miami-Dade County.

5. Merge the geo-allocated data to commuting zones and their characteristics using their 1990 county designations and the Equality of Opportunity data

Now, I describe steps 2 and 3 in more detail.

Step 2. Since DQ data is allocated to 2000 Census geographies, but HMDA LAR data is allocated to geographies of several vintages, each transaction needed to be paired with its Census geographies, up to the tract level, from 1990 and 2010. Furthermore, to establish very local demographic attributes of neighborhood, I actually performed this pairing up to the block group level. To do so, for each geography vintage, I performed an "st_within" merge from the "sf" R package; this checks each observation in the "left" dataset (the DQ transaction data) and finds all geographies that point is "within" from the "right" dataset (the 1990, 2000, and 2010 Census geographical data). The 2000 merge matched DQ designated Census tracts well over 99% of the time, suggesting a valid matching system; importantly, the 1990 and 2010 merges matched DQ designated Census tracts at a far lower rate, suggesting that this geographical processing was quite necessary. Then, a "HMDA tract" variable was created by selecting the Census tract from the geography vintage that was being used by HMDA LAR data at the time the transaction occurred (see above).

Step 3. This geo-allocated data was matched to HMDA LAR data using "HMDA tract," loan amount, and lending institution. Manually considering many pairs of records suggested that loan amounts either matched perfectly[2] or could not reasonably be said to likely reference the same transaction. Thus, I enforced perfect matching on tract and loan amount. It was immediately obvious, however, that the strin g variable identifying the lending institution matched perfectly only on extremely rare occasions,

[2] up to the thousands, which is the level at which the HMDA LAR records the variable

meaning a fuzzy merge would be necessary, though I did enforce perfect matching on the first letter. To prepare for a probabilistic matching algorithm, I altered the lender strings by converting to all uppercase, removing common words like "mortgage" and "company," and removing all spaces and punctuation.

The "RecordLinkage" R package implements several algorithms for performing a probabilistic merge, documented in (Sariyar and Borg, 2010); I chose the expectation maximization algorithm and the Jaro-Winkler string distance metric, which together yield a score for each possible pair that matches on the "mandatory match," or "blocking" variables. The final step was to find a cutoff score—a score above which a pair would be deemed a match and below which it would not. I performed this final step through manual clerical review; fortunately, it was quite simple to determine that an appropriate cutoff was approximately 0.95, as almost all pairs above that cutoff were a clear match, and almost all pairs below were a very poor match.

B.2 Further Details on Machine Learning Methods

In the text, I refer to two machine learning methods to flexibly control first for income alone, and later for a variety of demographic attributes of a neighborhood. Here, I provide more detail about the methods.

I use gradient boosted trees, implemented in the popular package "LightGBM" (`https://github.com/Microsoft/LightGBM`). This ensemble method builds successive regression trees, each trained to predict the error from the previous collection of trees; the final result for a given observation is the sum of the values of all trees applied to that observation.

For this method, no normalization of explanatory or dependent variables is necessary. All variables are kept at their nominal values, and the algorithm attempts to find the best relationship between them that it can. It is an extremely flexible method in that it imposes no functional form structure on the relationship between the variables, and it also allows for different relationships in different commuting zones while avoiding the overfitting that occurs through simple demeaning or fixed effects.

It is principally regularized to avoid overfitting by placing limits on the number of leaves in any given tree; choosing a learning rate—successive trees are in fact trained on only a fraction of the error from the previous collection; and stopping the algorithm after building a certain number of such trees are grown. These last two parameters should be loosely inversely related; if each tree accounts for a larger percentage of the previous error, one should need fewer of them. Assuming the number of trees is correctly adjusted, as will be discussed shortly, there is a tradeoff in manipulating the learning rate: a higher learning rate runs more quickly, but is less precise both in and out of sample.

I chose the number of leaves by fixing a high learning rate, and then using 3-fold cross validation on a variety of tree sizes. With tree sizes selected, I then used 3-fold cross validation again to choose the learning rate. During each iteration of the cross validation procedure, the number of trees was chosen by continuously testing for improvement in out of sample error on a 10% holdout sample; the number of trees used for the chosen hyperparameters was then used for subsequent training of the algorithm. 300 leaves and a learning rate of 0.01 was used for the prediction using only income, date, and commuting zone, while 500 leaves and a learning rate of 0.02 was used for the prediction using all variables in Figure 2.12. I did not separately cross-validate on each race-specific dataset used in Section 2.4.

Proper standard errors account for uncertainty in the predictions resulting from this algorithm, which I have calculated with an appropriately blocked bootstrap. I did not separately cross-validate for each bootstrap iteration.

Appendix C

Appendix to Chapter 3

C.1 Proof of Proposition 3.2.1

Individual i faces the following optimization problem:

$$V(w_i(1-\tau_j^L), h_j(1+\tau_j^H)) = \max_{C,H,L} \epsilon_{ij} \left\{ C^a H^b - \frac{\theta}{1+\gamma} L^{1+\gamma} \right\}$$

s.t.

$$C + Hh_j(1+\tau_j^H) = w_i(1-\tau_j^L)L$$

First consider the problem given a certain amount of labor supplied. Then, the individual must divide his income between housing and all other consumption. Since utility over these goods is Cobb-Douglas, the income shares are fixed, meaning that

$$C = \frac{a}{a+b} L w_i (1-\tau_j^L)$$
$$H = \frac{b}{a+b} \frac{L w_i (1-\tau_j^L)}{h_j(1+\tau_j^H)}$$

Now the individual faces an unconstrained optimization over the amount of labor to supply:

$$V(w_i(1-\tau_j^L), h_j(1+\tau_j^H)) = \max_L \epsilon_{ij} \left\{ \frac{a^a b^b}{(a+b)^{a+b}} [h_j(1+\tau_j^H)]^{-b} [L w_i (1-\tau_j^L)]^{a+b} \right.$$
$$\left. - \frac{\theta}{1+\gamma} L^{1+\gamma} \right\}$$

The first order condition is

$$\frac{a^a b^b}{(a+b)^{a+b}} [h_j(1+\tau_j^H)]^{-b} [L w_i(1-\tau_j^L)]^{a+b-1} = \theta L^\gamma$$

Rearranging yields

$$L(w(1-\tau_j^L), h_j(1+\tau_j^H)) = \tilde{L} \cdot (h_j(1+\tau_j^H))^{\frac{-b}{\gamma+1-a-b}} (w(1-\tau_j^L))^{\frac{a+b}{\gamma+1-a-b}}$$

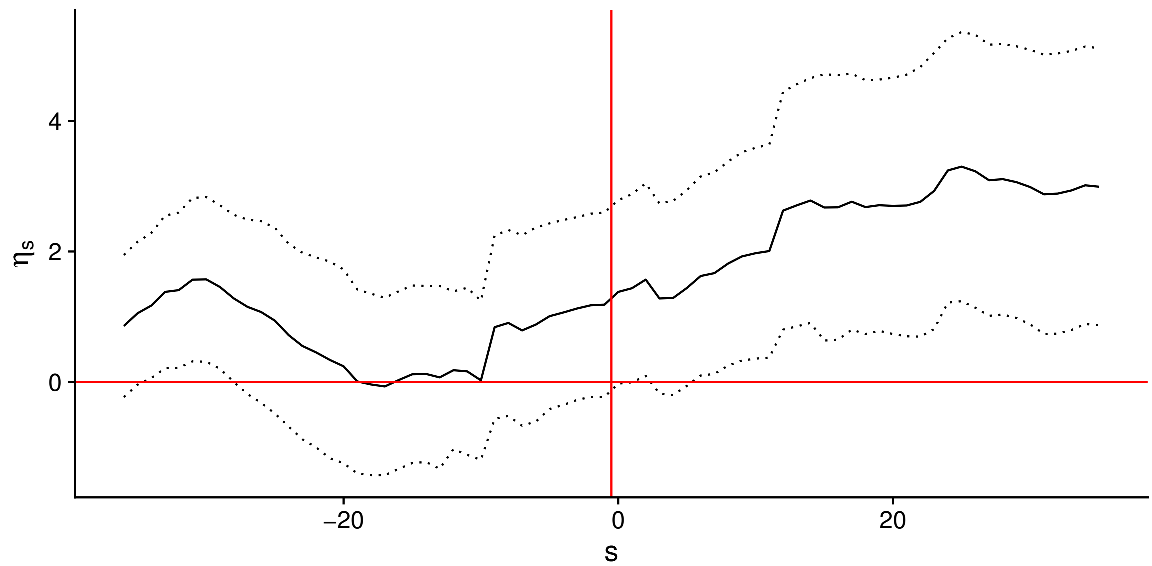
where

$$\tilde{L} = \left(\frac{a^a b^b}{(a+b)^{a+b}\theta} \right)^{\frac{1}{\gamma+1-a-b}}.$$

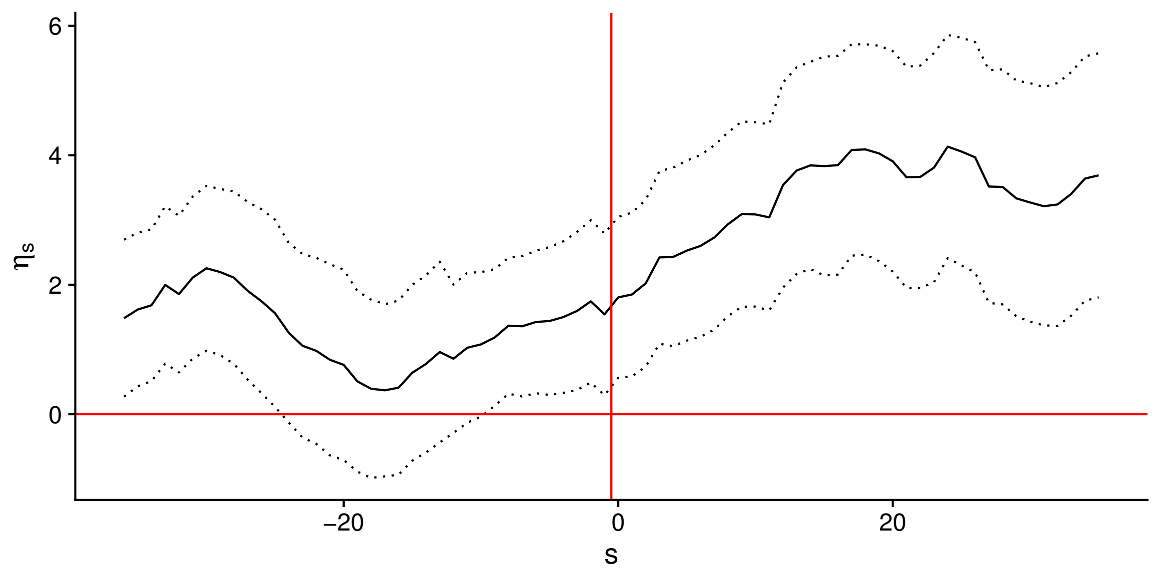
Substituting this into the expressions for C and H yields the stated forms, where $\tilde{C} = \frac{a}{a+b}\tilde{L}$ and $\tilde{H} = \frac{b}{a+b}\tilde{L}$. Substituting all of these into the utility function yields the proposed indirect utility function, where

$$\tilde{V} = \tilde{C}^a \tilde{H}^b - \frac{\theta}{1+\gamma} \tilde{L}^{1+\gamma}.$$

C.2 Dynamic Estimates for Other Samples

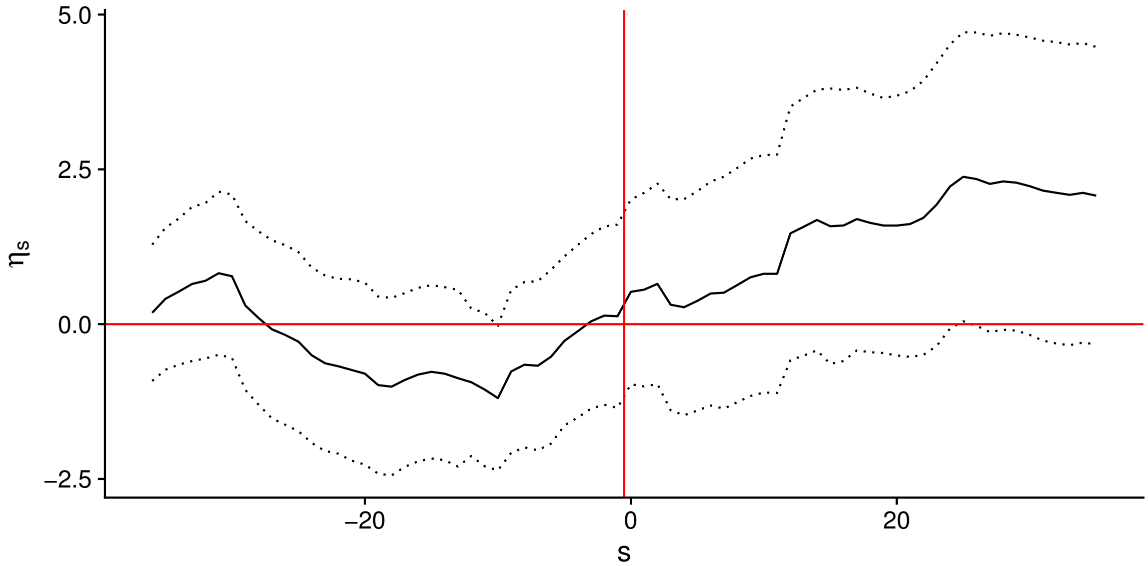


(a) Taxes on families earning \$50,000, median home prices

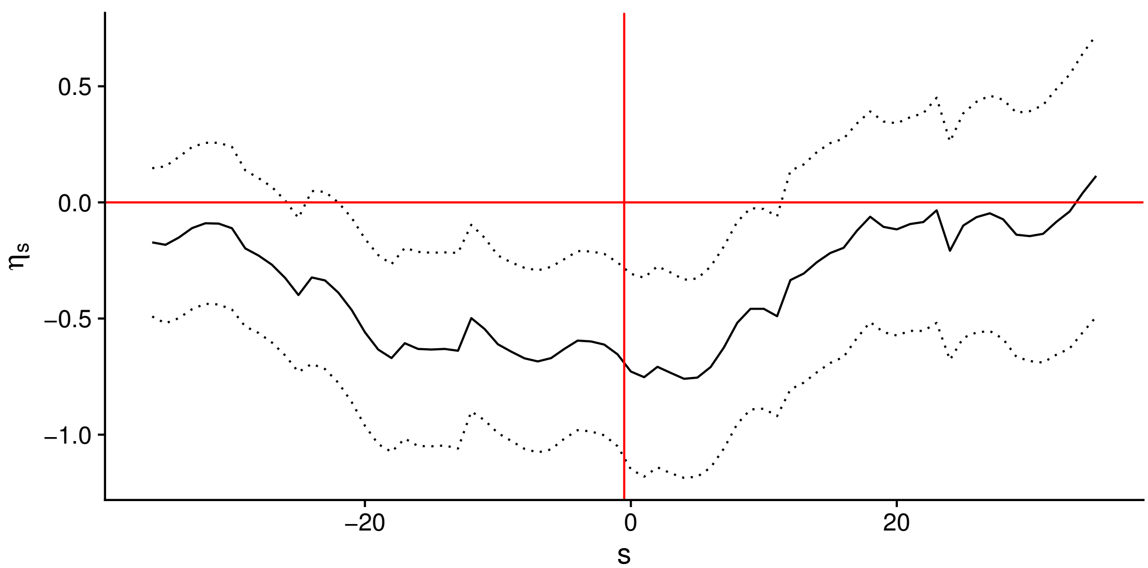


(b) Taxes on families earning \$50,000, per square foot home prices

Figure C.1: Dynamic price responses to tax changes via estimation of Equation 3.5. Elasticities of home prices in month $t + s$ with respect to a permanent change in the net-of-tax rate in month t are given by the solid line, with 95% confidence intervals bounded by the dotted lines. Standard errors are clustered separately at the border segment and state levels.



(c) Taxes on families earning \$100,000, median home prices



(d) Top marginal tax rates, median home prices

Figure C.1 (Continued): Dynamic price responses to tax changes via estimation of Equation 3.5. Elasticities of home prices in month $t + s$ with respect to a permanent change in the net-of-tax rate in month t are given by the solid line, with 95% confidence intervals bounded by the dotted lines. Standard errors are clustered separately at the border segment and state levels.