NBER WORKING PAPER SERIES

EQUILIBRIUM ANALYSIS IN THE BEHAVIORAL NEOCLASSICAL GROWTH MODEL

Daron Acemoglu
Martin Kaae Jensen

Working Paper 25363
http://www.nber.org/papers/w25363

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2018

We thank Xavier Gabaix for very useful discussion and comments. Thanks also to Drew Fudenberg, Marcus Hagedorn, David Laibson, Paul Milgrom and Kevin Reffett, as well as participants at the TUS-IV-2017 conference in Paris, and seminar participants at Lund University and the University of Oslo for helpful comments and suggestions. Acemoglu gratefully acknowledges financial support from the Army Research Office. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Equilibrium Analysis in the Behavioral Neoclassical Growth Model
Daron Acemoglu and Martin Kaae Jensen
NBER Working Paper No. 25363
December 2018
JEL No. D50,D90,O41

ABSTRACT

Rich behavioral biases, mistakes and limits on rational decision-making are often thought to make equilibrium analysis much more intractable. We show that this is not the case in the context of the neoclassical growth model (potentially incorporating incomplete markets and distortions). We break down the response of the economy to a change in the environment or policy into two parts: a direct response at a given vector of prices, and an equilibrium response that plays out as prices change. We refer to a change as a "local positive shock" if the direct response, when averaged across households, increases aggregate savings. Our main result shows that under weak regularity conditions, regardless of the details of behavioral preferences, mistakes and constraints on decision-making, the long-run equilibrium will involve a greater capital-labor ratio if and only if we start with a local positive shock. One implication of this result is that, from a qualitative point of view, behavioral biases matter for long-run equilibrium if and only if they change the direction of the direct response. We show that these aggregate predictions are coupled with individual-level "indeterminacy": nothing much can be said about individual behavior.

Daron Acemoglu
Department of Economics
MIT
50 Memorial Drive
Cambridge, MA 02142-1347
and CIFAR
and also NBER
daron@mit.edu

Martin Kaae Jensen
School of Economics
University of Surrey
United Kingdom
m.k.jensen@surrey.ac.uk

1 Introduction

Most standard macro and growth models rely on very restrictive behavioral assumptions about households — infinitely lived, often representative, agents who are capable of solving complex maximization problems without any behavioral biases or limitations, and of implementing the optimal decisions without any inconsistencies or mistakes. It is an uncomfortable stage of introductory graduate courses when these assumptions are introduced and students rightfully ask whether everything depends on them. A natural conjecture is that these assumptions do matter: not only do general equilibrium effects become notoriously complicated and the set of indirect effects correspondingly rich; we would also expect the specific departure from full rationality — *e.g.*, systematic mistakes, ambiguous beliefs, overoptimism or dynamic inconsistency — to have a first-order impact on how the economy responds to changes in policy or technology. In this paper, we show that robust results about the long-run response of a one-sector neoclassical economy to changes in policy or technology can nonetheless be obtained in the presence of general behavioral preferences.

Suppose, for example, we would like to analyze the implications of a reduction in the capital income tax rate on the long-run level of the capital stock. Starting from an initial steady-state equilibrium, we can break this analysis into two steps: first, we determine the *direct response*, measuring the impact of the policy change at the given vector of prices (determined by the initial capital-labor ratio). We refer to such a policy change as a "*local positive shock*" if the average of the direct responses of households leads to an increase in aggregate savings. Second, we have to determine the subsequent *equilibrium response*, which involves tracing the change in prices and the resulting change in household behavior and the capital stock necessary for the economy to settle into a new steady-state equilibrium. It is this second step that is generally challenging. To illustrate this, suppose that there are two groups of households. The first is responsive to the after-tax rate of return to capital and increases its savings. This raises the capital stock and wages, creating a negative income effect on savings. If the second group has a powerful income effect or behavioral biases that make it reduce its savings, its *equilibrium* (indirect) response might dominate the response of the first group, making it impossible to say anything about how the long-run capital stock will change.

Against this background, the current paper establishes that in the "behavioral neoclassical growth model" — meaning the one-sector neoclassical growth model but allowing for a rich set of consumer behaviors, heterogeneity, and uncertainty, as well as for incomplete markets and distortions — these equilibrium effects will never reverse the direct response.[1] So if the direct

[1]Note that here "neoclassical" only refers to the production side of the economy and does not presume or impose

response is a local positive shock, the long-run capital stock will necessarily increase; and if the direct response is a local negative shock, the long-run capital stock will decrease.[2] Notably, only minimal regularity conditions are imposed: the result remains valid under a rich set of behavioral biases and limitations on rational decision-making. Also noteworthy is that these strong predictions about aggregate behavior are true even though nothing general can be said about individual behavior — many groups of individuals, not just those that are making mistakes or are subject to severe behavioral biases, may react in the opposite way and reduce their savings in equilibrium. But there will always be sufficiently many other households who increase their savings for the economy's aggregate equilibrium response to move in the right direction (meaning in the same direction as the initial impulse).

The intuition for this result can be seen at two complementary levels. The first is economic in nature and it is related to an idea that already appears in Becker (1962) that "aggregation" disciplines economic behavior. Though we cannot say anything about individual behavior, we can determine the behavior of market-level variables (that is, aggregates such as the capital stock and income per capita). This is because even if many households respond in the opposite direction of the direct response, in equilibrium enough households have to move in the same direction as the direct response. The second intuition for our result is more mathematical. To develop this intuition, suppose that the steady-state equilibrium is unique, and focus on a local positive shock. This initial response then increases the capital stock, and the only way the new steady-state equilibrium could have lower capital stock is when the equilibrium response goes in the opposite direction and more than offsets the impact of this initial positive shock. This in turn can only be the case if a higher capital stock induces lower savings. But even if this were the case, the equilibrium response could not possibly overturn the direct response. This is because the economic force leading to lower savings would not be present if the new steady-state equilibrium ended up with a lower capital stock, and thus the indirect equilibrium response would in this case reinforce rather than overturn the direct effect of a positive shock. When there are multiple steady-state equilibria, this reasoning would not apply to all of them, but a similar argument can be developed for extremal (greatest and least) steady-state equilibria, and under multiplicity, it is these equilibria to which our conclusions apply.

---

any rationality requirements on households. Nor does it impose complete markets. For example, according to this terminology a version of the Ramsey-Cass-Koopmans model with dynamically inconsistent preferences (Laibson (1997)) and/or various distortions on the producer side is a behavioral neoclassical growth model, and so is the Aiyagari model (Aiyagari (1994)) with or without fully rational households.

[2] As we explain later (see in particular footnote 22), this direct response may or may not correspond to the observed impact effect of a shock because it needs to be evaluated at the "long-run beliefs" of households. When there is no uncertainty or when long-run and short-run beliefs coincide, the direct response is the same as the impact effect. Otherwise, the two differ.

To establish that these conclusions and intuitions hold under fairly general specifications of mistakes and behavioral assumptions, we develop a general framework that nests a rich set of behavioral models of consumption-saving decisions. We then go through several canonical models of behavioral deviations from infinite-horizon maximization and show that they satisfy the weak regularity conditions we require for our conclusions to apply. These include models with non-time-separable preferences, (quasi-)hyperbolic discounting, preferences featuring self-control and temptation problems, various models of complexity-constrained maximization, models of sparse maximization and models of mistakes and non-rational expectations (see references below).

It is useful to step back at this point and clarify what the message of the paper is. Beyond providing a general framework for obtaining (qualitative) comparative statics under a rich set of behavioral assumptions, the paper characterizes when, in the context of the behavioral neoclassical growth model, behavioral richness and biases matter. Our main result says that any biases that work through the equilibrium responses, while maintaining that the initial changes in the environment correspond to a local positive shock, do not matter for qualitative conclusions (though of course they may be quantitatively important). But conversely, our result also clarifies that any behavioral biases that determine whether a given change in policy or environment is a local positive or negative shock will matter greatly. For example, we illustrate in Section 5.4 that a shock such as a reduction in the capital income tax rate that is a local positive shock with forward-looking perfect maximizers may become a local negative shock for an economy that houses a fraction of biased agents. In this scenario, our main theorem applies in reverse, and shows that because behavioral biases have turned the initial change in environment into a negative shock, all equilibrium responses coming from rational behavior or markets will not be able to reverse this, and the impact on the long-run equilibrium will (robustly) be the exact opposite of what one might have expected with fully rational agents. We should also reiterate at this point that our results are on long-run (steady-state) responses and do not characterize the dynamics of an economy.

Our paper is related to several literatures. The first, already mentioned, is Becker (1962)'s seminal paper which argues that market demand curves will be downward sloping even if households are not rational because their budget constraints will put pressure for even random behavior to lead to lower demand for goods that have become more expensive. Machina (1982) makes a related type of observation about the independence axiom in expected utility theory. Though related to and inspired by these contributions, our main result is very different. While Becker's argument is about whether an increase in price will lead to a (partial equilibrium) change in aggregate behavior consistent with "rational behavior", our focus is about taking the initial change in behavior, whether or not it is rational, as given and then establishing that under general conditions on the objectives and behavioral biases and constraints of households the (general) equilibrium responses will not reverse this direct effect.

As our overview in the next section clarifies, the second literature we build on is robust comparative statics (Topkis (1978), Vives (1990), Milgrom and Shannon (1994), Milgrom and Roberts (1994), Milgrom (1994), Quah (2007)). Not only do we share these papers' focus on obtaining robust qualitative comparative static results, but we also use similar tools, in particular a version of the "curve-shifting" arguments of Milgrom and Roberts (1994) (see also Acemoglu and Jensen (2015)) which allow us to derive robust results in non-monotone economies.[3] Nevertheless, our main theorem is not an application of any result we are aware of; rather, it significantly extends and strengthens the approach used in the robust comparative statics literature. We provide a detailed technical discussion of the relationship of our results to the previous literature in Appendix A. Most significantly, the notion of local positive shock used here for deriving global comparative static results requires behavior to increase only at a *specific* capital-labor ratio (or vector of prices) rather than the much stronger notion that behavior increases everywhere imposed in this literature.[4] As a result, we are able to establish that any initial change that is a local positive shock — in the sense that the sum of the initial responses of all agents is positive at the initial capital-labor ratio — combined with weak regularity conditions leads to sharp comparative static results.[5]

In this context, it is also useful to compare our results to those of our earlier paper, Acemoglu and Jensen (2015), where we analyzed a related setup, but with three crucial differences. First, and most importantly, there we focused on forward-looking rational households, thus eschewing any analysis of behavioral biases and their impacts on equilibrium responses. Second, and as a result of the first difference, we did not have to deal with the more general problem considered here, which requires a different mathematical approach. Third, we imposed considerably stronger assumptions to ensure that the direct response of all households went in the same di-

[3]See p.590 in Acemoglu and Jensen (2015) for additional discussion of such non-monotone equilibrium comparative statics results.

[4]See for example Lemma 1 (and Figures 1-3) in Milgrom and Roberts (1994) or Definition 5 in Acemoglu and Jensen (2015). Milgrom and Roberts (1994) also use local assumptions, but just to derive local comparative statics results (see Figure 7 and the surrounding discussion); this is different from our results, which are global despite being based on local assumptions.

[5]The literature on mean field games is also related, especially since some papers in this literature, in particular Light and Weintraub (2018), investigate comparative statics. Nevertheless, like the majority of the papers in robust comparative statics, these works also focus on changes in environments that correspond to uniform and global changes (for example, Theorem 3 in Light and Weintraub (2018)). See also Ahn, Kaplan, Moll, Winberry and Wolf (2018) and Achdou, Han, Lasry, Lions and Moll (2018) for applications of related ideas to the analysis of Bewley-Aiyagari-style models.

rection at all prices, which we do not do in the current paper.

Finally, our paper is related to several recent works that incorporate rich behavioral biases and constraints into macro models. These include, among many others, Laibson (1997), Harris and Laibson (2001), Krusell and Smith (2003), Krusell, Kuruscu and Smith (2010), and Cao and Werning (2017) who study the dynamic and equilibrium implications of hyperbolic discounting (building on earlier work by Strotz (1956), and Phelps and Pollak (1968)). Particularly noteworthy in this context is Barro (1999) who shows that many of the implications of hyperbolic discounting embedded in a neoclassical growth model are similar to those of standard preferences, but this is in the context of a model with a representative household and does not contain any comparative static results for this or other classes of behavioral preferences, which are our main contribution. Gul and Pesendorfer (2001, 2004) and Fudenberg and Levine (2006, 2012) develop alternative approaches to temptation and self-control and their implications for dynamic behavior. Koopmans (1960), Epstein and Hynes (1983), Kreps and Porteus (1978), Lucas and Stokey (1984) and Epstein and Zin (1989, 1991) develop richer models of dynamic behavior with non-time-separable preferences, and Becker and Boyd (1997) and Backus, Routledge and Zin (2004) develop certain macroeconomic implications of such preferences. Gilboa (1987), Schmeidler (1989) and Gilboa and Schmeidler (1995) develop models of decision-making with max-min features resulting from lack of unique priors, and Hansen and Sargent (2001, 2010) and Hansen, Sargent and Tallarini (1999) discuss related preferences in various macroeconomic applications. Recent important work by Gabaix (2014, 2017) considers the macroeconomic implications of bounded rationality resulting from the inability of individuals to deal with complex problems and their need to reduce it to a sparse optimization problem, while Sims (2003) and Woodford (2013) consider the consequences of other complexity constraints on optimization. Finally, there are several examples of models featuring (systematic) mistakes and near-rational behavior including Simon (1956), Luce (1959), McFadden (1974), McKelvey and Palfrey (1995), and Train (2009). In the context of expectation formation and their implications for macroeconomics classic references include Cagan (1956), Nerlove (1958) and more recently Fuster, Herbert and Laibson (2012) and Beshears et al (2013). None of these papers develop comparative statics for macroeconomic models that apply under general behavioral preferences.

The rest of the paper is organized as follows. Section 2 provides an informal overview of our approach and main results. In Section 3, we describe our general setup and also present a number of behavioral dynamic consumption choice models that are covered by our results. Section 4 contains our main results. Section 5 investigates individual behavior. We provide sufficient (but strong) conditions under which certain changes in environment are local positive shocks and demonstrate that even though we have sharp results on aggregate behavior, generally very little can be said about individual behavior. Section 6 verifies that the assumptions we impose on individual behavior hold in many of the most popular behavioral models of limited rationality. Section 7 concludes, while Appendix A contains an abstract discussion of our comparative statics results and some additional results in this respect, and Appendix B contains omitted proofs from the text.

2 Overview of the Argument

The objective of this section is to provide a non-technical overview of our argument, which is helpful both to understand our main results and as a roadmap for the rest of the paper.

To motivate our main focus, suppose the government reduces the capital income tax rate in order to increase the capital stock and aggregate output in the long run. Such a policy may be expected to achieve this objective if both of the following are true: (1) the direct response to the policy at the initial capital-labor ratio goes in the right direction and increases aggregate savings; (2) as the economy adjusts to this initial impetus and prices change as a result of the responses of all of the households in the economy, this initial impact will not be undone. The first supposition is only about individual responses — since we are holding prices constant. This is what we summarize with the term *local positive shock*. Though sometimes determining whether a change in parameters or policy is a local positive shock may be far from trivial, economically this is a straightforward step because it involves no statement about *equilibrium behavior*, *i.e.*, about how the capital-labor ratio and individual behaviors adjust jointly and settle into a new equilibrium. On the other hand, even with forward-looking rational households, equilibrium responses are quite complex, for example because of countervailing income and substitution effects. They become much richer once we depart from the benchmark of forward-looking, perfectly rational decision-making. In fact, the general presumption in the literature is that this richness makes equilibrium analysis very difficult or impossible. Our main result stands in contrast to this presumption: under fairly weak regularity conditions, local positive shocks will always lead to an increase in the long-term capital stock in the context of a general class of neoclassical growth models, and thus once we are able to determine that a change in environment is a local positive shock, almost no additional work is necessary for determining the direction of change of the long-run equilibrium, even under very general behavioral preferences and biases.

To explain these ideas more clearly, let us now focus on the one-sector neoclassical growth model with exogenous labor supply (which we normalize to unity). Suppose that the per capita production function is $f(k)$ and satisfies all the standard assumptions where k as usual denotes the capital-labor ratio. As in the rest of our analysis, we allow distortions or taxes which the households also take as given, and thus the rental rate of return on capital is $R(k) = (1 - \tau(k))f'(k) - \delta$, and the wage rate is $w(k) = (1 - \omega(k))(f(k) - f'(k)k)$, where δ is the depreciation rate, and $\tau(k)$ and $\omega(k)$ denote the distortions that apply, respectively, to capital and labor. These distortions could result from taxes, contracting frictions or monopoly distortions. Each household takes these functions as given and we assume that they are real-valued and smooth. The benchmark model without distortions is obtained by setting them equal to zero.

The richness in our framework originates in the household side. We proceed in two steps: we first suppose that there is a representative household and no uncertainty, and then we consider the heterogenous agents setting under uncertainty.

2.1 The Main Result with a Representative Household

With a representative household (in a deterministic environment), the household side of the economy can be summarized by a consumption function $c_{w,R}(k)$ where the subscripts w and R designate the dependence of this function on the rate of return on capital and the wage (equivalently, we can begin with the savings function $s_{w,R}(k) = (1 + R)k + w - c_{w,R}(k)$). The derivation of this consumption function in the standard case with forward-looking fully rational agents is straightforward. It can also be characterized similarly, even if with more work, when preferences are non-standard, such as quasi-hyperbolic ones as in Laibson (1997) or non-additive ones as in Epstein and Hynes (1983), or when there are mistakes and additional constraints on rational decision-making. In Section 3.1 we consider a variety of underlying behavioral consumption and saving models, for example, incorporating limited attention or computational constraints or hand-to-mouth consumption decisions (such as when $c_{w,R}(k) = \alpha \cdot ((1 + R)k + w)$ where $\alpha \in (0, 1)$ is the constant average propensity to consume). In general, we allow consumption to be multivalued (a correspondence) and the only substantive assumption we make is that the consumption correspondence is upper hemi-continuous, and increases in the assets of the representative household less than one-for-one, so that corresponding savings are increasing in assets.[6] In Section 6, we confirm that these restrictions are weak and reasonable — a diverse set of preferences satisfy them. Here we focus on the simpler case with unique consumption decisions and a continuous consumption function.

We are now in a position to define a key object in our analysis, the *market correspondence*,

[6]Upper hemi-continuity in particular allows for the consumption function or the market correspondence to have jumps which is possible under some of the preferences we would like to nest, such as (quasi-)hyperbolic discounting (see Laibson (1997), p.452).

given by

$$\mathcal{M}(k) = f(k) + (1-\delta)k - G(k) - c_{w(k),R(k)}(k)\ , \tag{1}$$

where $G(k) = \tau(k)f'(k)k + \omega(k)(f(k) - f'(k)k)$ is government consumption or waste created from distortions (below, we allow part or all of tax revenues and spending on distortions to be rebated to households). A *steady-state equilibrium* (or *equilibrium* or *steady state* for short) naturally satisfies

$$\mathcal{M}(k) = k,$$

and the characterization of this steady-state equilibrium is depicted in Figure 1 as the intersection between the market correspondence (the solid curve) and the 45° line. For simplicity, we start here with the case in which the shape of this market correspondence is such that there exists a unique intersection, denoted by k^* in the figure.

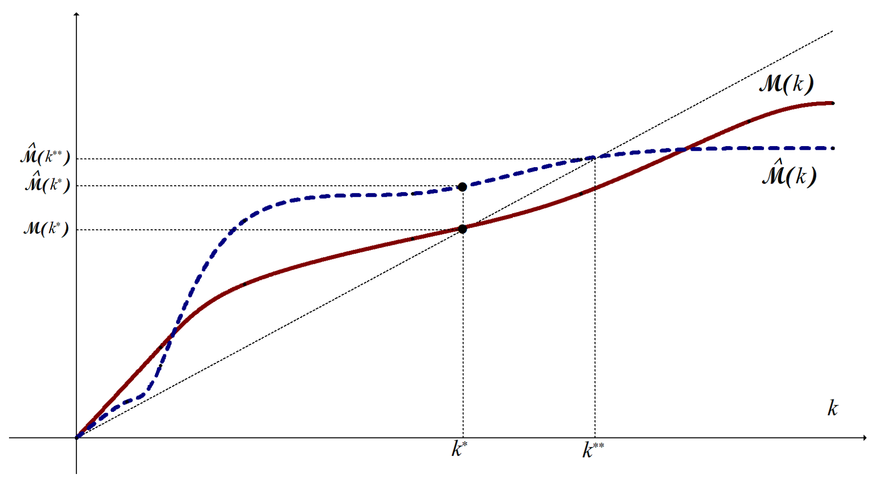

Figure 1: Market correspondences before and after a reduction in capital income taxes

Now with the help of this figure, we can clarify the discussion in the Introduction and provide an informal version of our main result. Let us first emphasize one feature of the market correspondence in Figure 1: the graph begins above and ends below the 45° line. Ending below the 45° line has the obvious meaning and follows directly from non-negativity of consumption and boundedness of feasible net output. As for beginning above the 45° line, this is automatically satisfied since $f(0) = G(0) = 0$, hence $0 = \mathcal{M}(0)$. The graph in Figure 1 lies (strictly) above the 45° line for k close to 0 which allows us to focus on the non-trivial steady state. This outcome is guaranteed under a standard Inada condition on f (at $k = 0$) if the average propensity to consume (APC) is bounded away from unity.[7]

[7] If the APC converged to 1 very rapidly, this could generate a "savings trap": the market correspondence would begin strictly below the 45° line and there would be multiple equilibria (unless the trivial equilibrium is the only one). We discuss multiplicity of equilibria below and in detail in Section 4.

Suppose that the economy starts at k^*. Consider a change in policy, for example a reduction in the capital income tax which shifts the function $\tau(k)$ down to $\hat{\tau}(k)$. Letting $\Delta(k) = (\tau(k) - \hat{\tau}(k))f'(k) > 0$ denote the change in the capital income tax for capital-labor ratio k, the policy thus increases the after-tax rate of return at capital-labor ratio k from $R(k)$ to $R(k) + \Delta(k)$. It is intuitive that such a cut in the capital income tax should encourage more savings at a given vector of factor prices or equivalently at the initial capital-labor ratio k^* — before any of the equilibrium responses kick in.[8] Mathematically, this amounts to

$$s_{w(k^*),R(k^*)+\Delta(k^*)}(k^*) \geq s_{w(k^*),R(k^*)}(k^*) \ . \tag{2}$$

To see how this relates to the notion of local positive shock discussed previously, note that since $-\Delta(k)k$ is the change in tax revenue, the market correspondence must change to the dashed curve in Figure 1, given by

$$\hat{\mathcal{M}}(k) = f(k) + (1 - \delta - \Delta(k))k - G(k) - c_{w(k),R(k)+\Delta(k)}(k) \ . \tag{3}$$

Crucially, if we evaluate (1) and (3) at $k = k^*$, we see that (2) will hold if and only if,

$$\hat{\mathcal{M}}(k^*) \geq \mathcal{M}(k^*) \ . \tag{4}$$

Equation (4) says that the market correspondence "shifts up" at k^* and is the definition of a *local positive shock* at k^* in this setting. Intuitively, it requires that the direct response to the change in the capital income tax is to increase savings at the initial capital-labor ratio k^*. Notably, we are *not* requiring that the market correspondence shifts up everywhere, and indeed in the figure, $\hat{\mathcal{M}}(k^*) > \mathcal{M}(k^*)$, but this inequality is reversed at other levels of the capital-labor ratio. The local nature of this condition critically implies that we do not need information about how prices change in order to determine whether the change in policy will be a local positive shock, since we are focusing only on behavior at the capital-labor ratio k^* (thus only on behavior for a given vector of prices). In particular, we do not need information about how such changes in prices affect the effective tax in the new equilibrium k^{**}.

Our main result traces the implications of changes in equilibrium prices following such a local positive shock. The main conclusion is that, as illustrated in Figure 1, the capital-labor ratio in the new steady state will necessarily greater than at the original steady state. This result is proved formally in Theorem 1 below and is informally summarized here.

[8] The statement that a local positive shock increases savings at the initial vector of prices needs to be qualified for the case in which the change in policy or parameters encapsulated in θ directly impacts these prices, for example, when it takes the form of a direct tax or a cap on the wage rate or the interest rate. This is the reason why we typically emphasize the effect of a change in policy (or parameters) at the initial capital-labor ratio rather than at the initial vector of prices.

Result 1 *Consider the market correspondence before a change in the environment (e.g., a reduction in the capital income tax rate), $\mathcal{M}$, changes to $\hat{\mathcal{M}}$, where $\hat{\mathcal{M}}(k^*) \geq \mathcal{M}(k^*)$, and suppose that the regularity conditions mentioned above are satisfied. Then the new (steady-state) equilibrium k^{**} satisfies $k^{**} \geq k^*$ if and only if the change in the environment is a local positive shock at k^*.*

Intuitively, given the regularity conditions, all we need to know is that the initial change in the environment or policy is a local positive shock at k^*. This can be seen geometrically in Figure 1. Even though the new market correspondence $\hat{\mathcal{M}}(k)$ may be below the one before the change, $\mathcal{M}(k)$, for many capital-labor ratios, the new steady-state equilibrium cannot fall below k^*. The intuition for this result was already discussed in the Introduction and will be provided in greater detail below.

How important are the assumptions made so far, in particular, the representative household assumption, the assumption that there is no uncertainty, and the restriction to a unique steady-state equilibrium? We next explain that these restrictions can be relaxed.

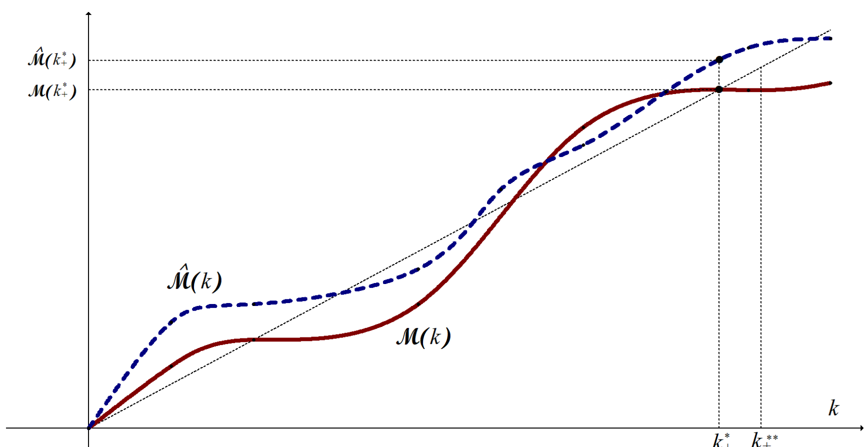

Figure 2: Extreme steady states satisfy the conclusion of Result 1, middle steady state does not.

Take first the assumption that there is a unique equilibrium. Figure 2 depicts a situation in which the market correspondence intersects with the 45° line multiple times. It is well known (*e.g.*, Milgrom and Roberts (1994)) that in this case middle equilibria may have perverse comparative statics, but we can establish similar results about extremal (greatest and least) equilibria. For example, the following is illustrated in Figure 2.

Result 2 *Suppose that the market correspondence satisfies the regularity conditions mentioned above. Let k^*_+ be the greatest equilibrium and consider a local positive shock at k^*_+ so that $\hat{\mathcal{M}}(k^*_+) \geq \mathcal{M}(k^*_+)$. Then the greatest equilibrium after the shock k^{**}_+ satisfies $k^{**}_+ \geq k^*_+$.*

2.2 Incorporating Uncertainty and Heterogeneity

Extending the previous observations to an environment with heterogenous households and uncertainty is conceptually straightforward, but necessitates a different mathematical approach, and developing such an approach is one of the main technical contributions of our paper. The added complication with heterogeneity comes from the fact that we can no longer work with a simple consumption function but need to take changes in the distribution of income into account. We show, however, that once the market correspondence is appropriately developed, the same insights hold. We now sketch this argument.

Suppose that prices and the aggregate capital stock are deterministic (no aggregate uncertainty) with households represented by the unit interval, $[0, 1]$. Once again take the capital-labor ratio $k > 0$ as given, and additionally fix an *asset distribution*, which is a measurable mapping $\lambda : (k, i) \mapsto \lambda^i$ that assigns a (possibly random) asset level λ^i to each household i in such a way that $\int_0^1 \lambda^i \, di = k$. This formulation is general enough to nest both the case in which there is a deterministic distribution of assets and/or preferences and the case where consumption decisions are random.[9] For a given asset distribution λ, we then define

$$\mathcal{M}_\lambda(k) = f(k) + (1 - \delta)k - G(k) - \int c^i_{w(k),R(k)}(\lambda^i) \, di. \tag{5}$$

Note that (5) is no more than an "accounting identity" (we are not at this point determining or restricting the asset distribution λ). What makes the definition useful is the next result which shows that by considering a suitably chosen *set* of asset distributions, we get a correspondence that gives us steady-state equilibria as fixed points and inherits the qualitative features of the simple case in (3).[10]

Result 3 *There exists a set of asset distributions Λ such that k^* is a (steady-state) equilibrium if and only if $\mathcal{M}_\lambda(k^*) = k^*$ for some $\lambda \in \Lambda$. Furthermore, under the regularity conditions imposed above, the market correspondence*

$$\mathcal{M}(k) = \{\mathcal{M}_\lambda(k) : \lambda \in \Lambda\} \tag{6}$$

is convex-valued and upper hemi-continuous, and its graph begins above and end below the $45°$ line. An equilibrium in this case is defined as

$$k \in \mathcal{M}(k).$$

[9] We discuss the technical details involved in defining the integral in the text (or in applying the appropriate law of large numbers). Note also that when λ^i is a random variable, $c^{\theta,i}_{w(k),R(k)}(\lambda^{\theta,i}(k))$ will also be a random variable, even if consumption is deterministic. When consumption is itself random, $c^{\theta,i}_{w(k),R(k)}(\lambda^{\theta,i}(k))$ will be a random variable even if λ^i is not. In either case (or when both apply), we again need to use a law of large numbers when defining the integral.

[10] For further details and proof, see Section 3.3. Specifically, the market correspondence (6) only coincides with average savings in equilibrium (see the discussion prior to Lemma 1).

This result is an informal version of our key lemma, Lemma 1, upon which the rest of our analysis builds. Results 1 and 2 generalize to environments with heterogeneity and uncertainty using this foundation. The main implication is that even though we do not know the equilibrium distribution of income/assets, the situation is conceptually no different from the representative household case. It is then straightforward to see graphically that analogues of Results 1 and 2 with heterogeneity and uncertainty hold once the market correspondence construction of Result 3 is used. This then is what allows us to establish the main message of this paper for a rich class of models featuring behavioral biases, mistakes, and other limits on rational behavior (see Sections 3 and 6).

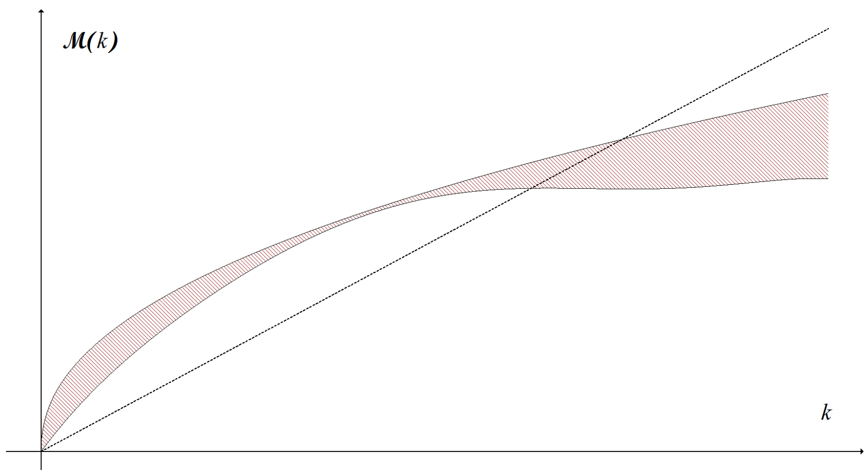

Figure 3: A market correspondence with the properties described in Result 3.

Two more observations are worth making. First, as we saw above, in the case with a representative household, a change in the environment is a local positive shock if the representative household increases its savings at the initial capital-labor ratio. A natural generalization to the case with heterogeneous households may appear to be to require that *all* households increase savings. Indeed, this is the type of assumption one typically adopts in the analysis of supermodular economies or games. However, what we need is much weaker than this, and a shift up of the market correspondence $\mathcal{M}(k)$ at k^* could result from some households increasing their savings while a significant fraction change their behavior in the opposite direction (see Section 5). Second, with a representative household, the fact that the equilibrium capital-labor ratio increases implies that the household has raised its savings in response to the change in the environment. With heterogeneity, there is no analogue of this result. In fact, as we show in Section 5.3, nothing can be said about how individual behavior changes in general. All we know is that aggregate savings increase so that the equilibrium capital-labor ratio increases, but this could be accompanied by a complex set of countervailing responses from some households in the economy. That our results hold without any implications at the individual level explains why we are able to obtain results without imposing rigid behavioral assumptions on households.

3 The Behavioral Neoclassical Growth Model and the Market Correspondence

This section presents our general model, formally defines the market correspondence, introduces Assumption 1, which encapsulates the key restrictions we impose on household behavior, and proves our key lemma (which establishes Result 3 from the previous section). Before moving to these key building blocks of our analysis, we start with a quick look at some of the behavioral models we incorporate in our general model. Our aim here is to provide high-level summary the types of behaviors our analysis focuses on. That these behavioral models satisfy our key Assumption 1 is established in Section 6.

3.1 Behavioral Consumption and Saving Models

The following examples are meant to briefly introduce some of the types of behaviors our general message applies to. We treat these, as well as a number of other models (for example Epstein and Zin (1989) type objectives, and quantal-response equilibrium) formally in Section 6. There, we also show that they all satisfy Assumption 1 below, and so by Lemma 1 lead to market correspondences that are qualitatively similar to the illustration in Figure 3. In all cases, this is also true with heterogenous households and any mix of the behaviors described next (*e.g.*, in Sections 5.3-5.4 we consider situations where some agents are "rational" while others follow rules-of-thumb).

3.1.1 Hyperbolic, Quasi-Hyperbolic, and General Delay Discounting

Consider the general delay discounted additive utility objective $U(c_0, c_1, c_2, \ldots) = u(c_0) + f(1)u(c_1) + f(2)u(c_2) + \ldots$. As shown by Strotz (1956), the only case where a household at date $t+1$ will necessarily wish to consume/save what it planned to consume/save at date t, is when $f(t) = \delta^t$ (geometric discounting). In all other cases, the objective will be dynamically inconsistent (and first-best behavior will be time-inconsistent, see Strotz (1956), Phelps and Pollak (1968), Loewenstein and Prelec (1992), Laibson (1997)). In such situations, the standard approach is to model the decision as a game between a sequence of temporal selves. We show in Example 1 below that such behavioral models fit into our general framework.

3.1.2 Random Utility, Mistakes, and Approximate Rational and Satisficing Behavior

Consider as in the previous model an additive objective but assume now that utility at each date is random: $U(c_0, c_1, c_2, \ldots) = u^{\epsilon_0}(c_0) + f(1)u^{\epsilon_1}(c_1) + f(2)u^{\epsilon_2}(c_2) + \ldots$. The random variable ϵ_t is interpreted as the household's idiosyncratic tastes/biases (McFadden (1974), p.108). There are two (mathematically equivalent) interpretations. The first is that the agent is uncertain about his future preferences, and if the objective is dynamically inconsistent, he is consequently uncertain about the behavior of future selves (see the previous example).[11] In the second interpretation, u^0 is a temporal self's true objective, and if $\epsilon_t \neq 0$, the agent consequently makes a mistake and maximizes an objective that departs from this true objective. In either situation, the agent's savings function — and when relevant, the savings function of future selves — will be a behavioral process (Train (2009), p.3). In the second interpretation, this behavioral process describes approximate rational behavior in the sense of Luce (1959). If the distribution ϵ_t is uniform on $[-a, a]$, $a > 0$, this can also be interpreted as satisficing/ϵ-optimizing behavior in the sense of Simon (1956).[12]

3.1.3 Sparse Maximization and Inattention

An individual faced with an infinite (or even just a long) time horizon may, optimally or as a rule-of-thumb, opt to keep down mental costs involved in estimating, assessing and using objective probabilities and calculating the optimal decision (Sims (2003)). One way to capture this in dynamic consumption and saving problems is to take as objective $\sum_{t=1}^{T} \beta^t u(c_t)$ where T is finite; so that the agent looks only T periods into the future at any point in time. Since at any future date, he will also look T periods into the future, such preferences are dynamically inconsistent; and the current self will thus take as given the expected (inattentive) behavior of future selves. This may be interpreted as a simple version of "sparse maximization" in the sense of Gabaix (2014, 2017). It can be combined with the random utility model above by taking the the time-horizon of future selves as an idiosyncratic characteristic of the household, so that the maximization problem becomes $\sum_{t=1}^{\epsilon} \beta^t u(c_t)$, where $\epsilon \in \{1, 2, \ldots, \hat{T}\}$ and the probability distribution over ϵ reflects the household's (subjective) beliefs about future selves' time-horizon. Here, the sparsity of the planning horizon at future dates is uncertain from the point of view of today, and the agent is uncertain about how inattentive/sparse future selves will be. Other, richer types of sparsity constraints following Gabaix (2014, 2017) can also be incorporated into this framework, for example, by reducing the set of choice variables, restricting the dependence of consumption and saving

[11] The game between temporal selves will in this situation be a Bayesian game.

[12] Whether an agent maximizes a function that is ϵ away from the true objective or ϵ-maximizes the true objective amounts to the same as long as the decision function is continuous in ϵ.

decisions on the states of nature, or more generally considering "sparse maximization" as we discuss further below.

3.1.4 Systematically Wrong Beliefs

In dynamic consumption and saving models, current consumption and savings depend on households' beliefs about future prices and parameters. If these beliefs are systematically wrong (*i.e.*, wrong period after period), any perfect commitment solution conditioned on these beliefs will not satisfy the budget constraints of the household, and is therefore inadmissible. The obvious solution is again to deal with the resulting dynamic inconsistency (which is now embedded in the belief structure) by modeling the dynamic consumption choice problem as a game between temporal selves. We show in Section 6 (Example 3) that this type of departure from perfectly rational behavior is covered by our setup as well.

3.1.5 Ambiguity

It is natural that households' beliefs/expectations, even if rational, are based on incomplete information about the objective probabilities governing any random disturbances. If the axioms of Savage (1954) are not satisfied, these (subjective) beliefs will not be uniquely determined (note that this has nothing to do with whether the beliefs are right or wrong). As we discuss in Example 5 in Section 6, most models of ambiguity are covered by our setup as well. In particular, agents/households may entertain multiple beliefs (Gilboa (1987), Schmeidler (1989), Gilboa and Schmeidler (1995)).

3.1.6 Rules-of-Thumb

Since our starting point below is savings and consumption functions, simple decision rules without any micro-foundation (rules-of-thumb) fit into the framework as well provided that they satisfy Assumption 1 below (in this context, this assumption is quite weak). For example, an agent might at any date simply save a fraction of current income with that fraction depending on some current variables such as a measure of the environment's variability. Just like the systematically wrong beliefs of Section 3.1.4, rules-of-thumb may include "highly irrational" behaviors.

3.2 Markets and Production

The production side is the same as the canonical neoclassical growth model (e.g., Acemoglu (2009)) augmented with general distortions.

Labor is in fixed supply and normalized to unity so we can use capital, capital-labor ratio and capital-per-worker interchangeably and denote it by k. Markets clear at all times, and production is described by a profit maximizing aggregate constant returns firm with a smooth (per capita) production technology $y = f(k)$ that satisfies $f(0) = 0$, $f' > 0$, and $f'' < 0$. We also impose that there exists $\bar{k} > 0$ such that $f(k) < k$ all $k \geq \bar{k}$, which ensures compactness. This condition is implied by the standard Inada conditions when these are imposed. The rate of depreciation is $\delta \in [0, 1]$.

As explained already in Section 2, our description allows for taxes and distortions $\omega(k)$ and $\tau(k)$ on labor and capital, and the wage and interest rate are therefore

$$w(k_t) \equiv (1 - \omega(k_t))(f(k_t) - f'(k_t)k_t) , \tag{7}$$

and

$$R(k_t) \equiv (1 - \tau(k_t))f'(k_t) - \delta . \tag{8}$$

The simplest example of such a distortion is a proportional tax, $\tau(k_t) = \tau$ on capital income and $\omega(k_t) = \omega$. Other examples include distortions from contracting frictions or markups due to imperfect competition. When $\tau(k) = \omega(k) = 0$ for all k, we recover the benchmark case with no distortions.

We allow proceeds from these distortions to be partially rebated to households (which will be the case when they represent taxes and some of the tax revenues are redistributed the households or when they result from markups that generate profits). The total amount of resources that is *not* rebated back to households (hence is either wasted or consumed by the government) is denoted by

$$G = G(k_t) . \tag{9}$$

If nothing is rebated back to households, then

$$G(k_t) = \omega(k_t)(f(k_t) - f'(k_t)k_t) + \tau(k_t)f'(k_t) . \tag{10}$$

On the other hand, if the only source of distortions is taxes and the government rebates everything back to consumers (*e.g.*, in the form of lump-sum transfers), then $G(k_t) = 0$.

3.3 Households and the Market Correspondence

We have already provided in Sections 2 and 3.1 some examples of the set of behaviors we would like to incorporate on the household side. We now formalize this by developing an abstract representation of household behavior (consumption/saving decisions) and then impose an assumption directly on this behavior (Assumption 1). We argue that this assumption is not very restrictive and also quite natural. We show in Section 6 that all of the behavioral household preferences outlined in Sections 2 and 3.1, as well as several others, satisfy it.

There is a continuum of households $[0,1]$ with a typical household denoted by $i \in [0,1]$. Any randomness is such that there is no aggregate uncertainty so capital k_t is deterministic at each date and factor prices are therefore given by (7) and (8).

The key object in our approach is the consumption/savings correspondence of households. Our focus on "correspondences" is motivated by our desire not to assume uniqueness, since this would rule out many of the behaviors discussed in Section 2 (for example, quasi-hyperbolic discounting, which typically leads to such non-uniqueness, see *e.g.* Laibson (1997), p.452). Household $i \in [0,1]$'s saving decisions will depend on:

1. its current asset level $a^i \in A^i \subseteq \mathbb{R}$;

2. after-tax interest and wage rates, R and w, which are assumed to be constant over time (the fact that these are constant is a consequence of our focus on steady states);

3. a random disturbance $z^i \in Z^i$ (where $Z^i \subseteq \mathbb{R}^m$), which represents any idiosyncratic randomness in the labor endowment or interest rate processes, or in preferences;

4. the households' beliefs P^i about future variables (conditional on their current observations and the current value of z^i). We are currently leaving what exactly beliefs incorporate unspecified, because it will vary from model to model but typically households' beliefs will include expectations about the stochastic processes of income, policy variables and factor prices R and w.

We denote by θ^M the "true model", which encapsulates all relevant information on the stochastic process that governs $z_t = (z_t^i)_{i \in [0,1]}$, actual policies, features of the production technology, market clearing conditions, and so forth (in brief, it includes everything in this section). We do not impose that households actually observe or use all of this information when forming beliefs; in the case of adaptive expectations, for example, they use none of it (see Section 6 which contains a detailed treatment of beliefs in the context of an explicit model of individual behavior).

We assume throughout that $(z_t^i)_{i \in [0,1]}$ is a Markov process with a unique invariant (ergodic) distribution μ_z.[13] Crucially, due to our long-run (steady state) focus, we also assume that beliefs P^i are time-invariant. Since, as noted above, P^i is conditioned on the current value of z^i, this assumption does not mean that individual beliefs are time-invariant, but simply that the mapping

[13]As in Acemoglu and Jensen (2015), we can allow for multiple invariant distributions and then deal with the multiplicity of steady states that results once again focusing on the results for the greatest and least equilibria.

from current values of z^i (and possibly from other current observations) into beliefs and expectations is stationary. Nevertheless, it does impose one important restriction: We are only focusing on models in which there exist well-defined ergodic or "long-run" expectations as represented by P^i.[14] This restriction is naturally satisfied in models without uncertainty, with "dogmatic beliefs" (for example, as discussed in Section 3.1.4 and in Example 3 below), in models with rational expectations as well as models with well-defined learning rules. It rules out models in which the beliefs of a positive measure of the households fail to converge to an ergodic distribution, P^i ($= \lim_{t\to\infty} P^i_t$). We return to this issue in Section 4.

A simple and familiar example is the Aiyagari model (Aiyagari (1994)), where z^i are *i.i.d.* labor endowment shocks $z^i \sim \mu_{z^i}$, and agents have rational expectations so beliefs about future prices coincide with actual (equilibrium) prices and beliefs about the future realizations of the labor endowment shock coincides with the objective probability measure, μ_{z^i}. For this reason, as in models with rational expectations more generally, beliefs can be suppressed/ignored altogether. In general we do not assume that z^i are independent across households as long as any dependency is consistent with the absence of aggregate uncertainty (the simplest case is when beliefs are independent conditioned on prices and policy). Beliefs P^i will clearly not be independent since they depend on the same set of information (prices, the model, etc.).

As a shorthand, we define an "*environment*", denoted by $\theta = (\theta^M, (P^i)_{i\in[0,1]})$, to summarize the true model θ^M and beliefs $(P^i)_{i\in[0,1]}$. Given an environment θ, we can then define the *saving correspondence* of household i as $S^{\theta,i}_{w,R,z^i}(a^i) \subseteq A^i$, which maps current variables, the model and the current asset level of the household, a^i, into a set of feasible asset levels for next period (or into a set of "gross savings"). We now state our main assumption on these saving correspondences, and return in Section 6 to verifying that this assumption is satisfied for the class of preferences we study in this paper.

Assumption 1 *For each $i \in [0,1]$, the savings correspondence $S^{\theta,i}_{w,R,z^i}(a^i)$ has compact range, and is upper hemi-continuous in w, R, and a^i and measurable in z^i, and is increasing in a^i.*

Several points of clarification are useful at this point. First, a correspondence is *measurable* if the inverse image of any open set is Borel-measurable (Aubin and Frankowska (1990), p.307). Second, the savings correspondence $S^{\theta,i}_{w,R,z^i}(a^i)$ is *increasing* in a^i if and only if it is ascending in the sense of Topkis (1978), or more explicitly, $S^{\theta,i}_{w,R,z^i}(a^i)$ is increasing if its greatest and

[14]Since savings decisions are not assumed to be unique, households may have multiple beliefs (for Assumption 1 below to hold, the set of such beliefs must be (weak*-) upper hemi-continuous in w, R, and z and have a compact range). Multiple beliefs arise if we (as modelers) are unable to discriminate between different possible household beliefs (footnote 18). Note also that beliefs may be non-additive as described in Example 3.1.5 above and in Section 6.

least selections are (weakly) increasing in a^i. In words, next period's assets lie within an interval that weakly increases with current assets (when S^i is single-valued, it simply means that next period's asset holdings do not decrease if the current period's asset holdings increases). Third, the savings correspondence $S^{\theta,i}_{w,R,z^i}(a^i)$ is said to have a *compact range* if for fixed w and R, $S^{\theta,i}_{w,R,z^i}(a^i) \subseteq \bar{A}^i$, all z^i and a^i for some compact subset $\bar{A}^i \subset \mathbb{R}$ (note that $\bar{A}^i$ may depend on w and R so it is possible for households' savings to go to infinity as prices go to 0 or infinity). A compact range ensures that for fixed prices, the households never accumulate infinite assets or amass arbitrary debts both of which would present problems for existence of steady states. One normally derives the lower bound on accumulation from more fundamental transversality conditions or borrowing constraints and the upper bound by bounding the set of feasible consumption/savings sequences using effective compactness (Section 3.2). Imposing it directly is without much loss of generality in this setting. Finally, note that when S^i is upper hemi-continuous and has a compact range, greatest and least selections/savings functions always exist.

In what follows, instead of the saving correspondence $S^{\theta,i}_{w,R,z^i}(a^i)$, it turns out to be more convenient to work with the *induced saving distribution*, denoted by $S^{\theta,i}_{w,R}(a^i)$. Mathematically, this corresponds to the probability distribution of gross savings of a household, interpreted as random variables on A^i, induced by the stochastic saving decisions of the household. The induced saving distribution captures in a succinct way both the random nature of savings resulting from shocks z^i impacting the household (as in Aiyagari (1994) model, for example) and any stochastic elements resulting from the fact that a^i itself is a random variable or from additional randomization or indeterminacy in the household's behavior.[15] Using this induced saving distribution and a similarly defined induced consumption distribution $C^i_{w,R}$, we now proceed to introduce our notion of equilibrium and the key objects for our equilibrium analysis, the market correspondence.[16]

In the next definition, we define a state-state equilibrium in terms of the corresponding capital-labor ratio and the factor prices are then derived from this capital-labor ratio.

Definition 1 ***(Equilibrium)*** *The capital-labor ratio $k^* \in \mathbb{R}_+$ represents a (steady-state) equilibrium*

[15]Formally, let $Q^{\theta,i}(a^i, B) = \int_{Z^i} 1_{S^{\theta,i}_{w,R,z^i}(a^i)}(B)\mu_{z^i}(dz^i)$ be the transition correspondence of savings (where μ_{z^i} denote the marginal distribution of the invariant distribution of $z_t = (z^i_t)_{i\in[0,1]}$). For a random variable $\hat{a}^i$ on A^i with distribution η^i_t, we can now define $S^{\theta,i}_{w,R}(\hat{a}^i)$ as the set of random variables on A^i with distributions, $\eta^{\theta,i}_{t+1}(B) = \int_{a^i\in B} Q^{\theta,i}(a^i, B)\eta^{\theta,i}_t(da^i)$. Thus $S^{\theta,i}_{w,R}(\hat{a}^i)$ is the adjoint Markov correspondence (or rather, the set of random variable with distributions given by the adjoint; see the Appendix in Acemoglu and Jensen (2015) for more details).

[16]Alternatively, we could work in distributional strategies but the current approach is more natural in the macroeconomic context.

Note also that when the shock/random disturbance is in a stationary state, we may — even if savings and the random disturbance are in general correlated — "disintegrate" this stationary distribution to get the marginal distribution of savings as described (see the Appendix of Acemoglu and Jensen (2015) for details).

if equilibrium prices $w^* = w(k^*)$ and $R^* = R(k^*)$ are given by (7) and (8), the gross savings (assets) of household i is given by (the random variable) $\hat{a}^{*,i} \in S^{\theta,i}_{w^*,R^*}(\hat{a}^{*,i})$ for almost every $i \in [0,1]$, and the capital market clears, that is, $k^* = \int \hat{a}^{*,i}\, di$.

In this definition we are implicitly assuming that the integral $\int \hat{a}^{*,i}\, di$ is well-defined from some version of the law of large numbers.[17] Note also that the set of equilibria explicitly depends on the environment θ, and so do the functions designating the relationship between factor prices and the capital-labor ratio, $w(k)$ and $R(k)$ (for example, because θ includes taxes on factor payments or direct regulations on prices). Nevertheless, to simplify the notation we suppress the dependence of these functions on θ when this will cause no confusion. We return to the comparative statics of equilibria when the environment θ changes in Section 4.

We are now ready to formally define the market correspondence which was introduced in Section 2.

Definition 2 (Market Distributions and the Market Correspondence) *Let $C^{\theta,i}$ denote the consumption correspondence and $S^{\theta,i}$ denote the savings correspondence of household $i \in [0,1]$. Also, let $G(k)$ denote government consumption and distortionary waste given the capital-labor ratio k.*

- *A measurable mapping $\lambda : (i,k) \mapsto \lambda^i(k)$ where $\lambda^i(k)$ is a random variable on $A^i \subseteq \mathbb{R}$ is a market distribution, if*
$$\lambda^i(k) = \frac{\hat{a}^i(k)}{\int \hat{a}^i(k)\, di} k, \text{ for all } (i,k) \tag{11}$$
where $(\hat{a}^i(k))_{i\in[0,1]}$ solve the fixed point problem,
$$\hat{a}^i(k) \in S^{\theta,i}_{w(k),R(k)}\Big(\frac{\hat{a}^i(k)}{\int \hat{a}^i(k)\, di} k\Big) \, , i \in [0,1] \, . \tag{12}$$

- *The market correspondence $\mathcal{M}^\theta : \mathbb{R} \to 2^{\mathbb{R}}$ is*
$$\mathcal{M}^\theta(k) = \{Af(k) + (1-\delta)k - G(k)\} -$$
$$\{c \in \int C^{\theta,i}_{w(k),R(k)}(\lambda^i(k))\, di : \ \lambda^i(k) \text{ is a market distribution}\} \, . \tag{13}$$

[17]There is a large literature on laws of large numbers and their application in the presence of continuum of random variables as in our economy (Al-Najjar (2004), Uhlig (1996), Sun (2006)). Here and everywhere else in this paper we remain agnostic about precisely which formulation of the law of large numbers has been applied in the background. This "agnostic" approach is also the one taken in Acemoglu and Jensen (2015) where $\int a^i(k)\, di$ is simply *assumed* to equal (or be one-to-one) with a real number. This approach has the advantage of not committing to a specific interpretation and therefore comes with maximum generality. On the downside, we must be careful to not push the generality of the setting too far: In the Aiyagari model, for example, any sensible application of a law of large numbers will require that the labor endowments' conditional distributions are at least pairwise independent conditioned on k. For further details and references, see Acemoglu and Jensen (2010, 2015)).

Note that $\lambda^i(k)$ may be correlated across households (this will happen, for example, if households are subject to correlated shocks). But conditional on k, the definition of $\mathcal{M}$ requires that the integral $\int C^{\theta,i}_{w(k),R(k)}(\lambda^i)\ di$ has a degenerate distribution ($\int C^{\theta,i}_{w(k),R(k)}(\lambda^i)\ di$ interchangeably denotes both this distribution and its point of unit mass). With a representative household with consumption correspondence C, (11) reduces to $\lambda^i = k$ for all i, (12) becomes redundant, and (13) collapses to $\mathcal{M}^\theta(k) = Af(k) + (1-\delta)k - \int C^\theta_{w(k),R(k)}(k)\ di$. If, in addition, $\mathcal{C}^\theta$ is single-valued and we suppress θ, this brings us back to (1).

We are now ready to state and prove a formal version of Result 3 from the previous section which enables us to analyze models with rich heterogeneity in terms of behavior and preferences in a tractable manner. In particular, the next lemma establishes that we can work directly with the market correspondence defined as in Definition 2 (without specifying the exact distribution λ) and fixed points of the market correspondence will be steady-state equilibria. The proof of this lemma uses the Fixed Point Comparative Statics Theorem of Acemoglu and Jensen (2015) (Theorem 4, p.601, which itself builds on Smithson's generalized fixed point theorem) as well as Richter's Theorem (Aumann (1965)), but the most critical component is the observation that for a given k, $\mathcal{M}^\theta(k)$ equals the set of fixed points of a convex valued correspondence whose least and greatest selections are decreasing, and therefore it is itself convex-valued (see also the discussion immediately after the proof).

Lemma 1 *If all households satisfy Assumption 1, the market correspondence $\mathcal{M}^\theta$ is a compact- and convex-valued upper hemi-continuous correspondence that begins above and ends below the 45° line. Furthermore, $k \in \mathcal{M}^\theta(k)$ if and only if k is a steady-state equilibrium.*

Proof. Since $Af(k) + (1-\delta)k - G(k)$ equals aggregate income after taxes and net of any waste,

$$\mathcal{M}^\theta(k) = Af(k) + (1-\delta)k - G(k) - \int ((1+R(k))a^i + l^i w(k) - S^{\theta,i}_{w(k),R(k)}(\frac{\hat{a}^i(k)}{\int \hat{a}^{\theta,i}(k)\ di}k)\ di$$

$$= \int S^{\theta,i}_{w(k),R(k)}(\frac{\hat{a}^i(k)}{\int \hat{a}^i(k)\ di}k)\ di\ .$$

Hence

$$\mathcal{M}^\theta(k) = \{\int \hat{a}^i(k)\ di :\ \hat{a}^i(k) \in S^{\theta,i}_{w(k),R(k)}(\tfrac{\hat{a}^i(k)}{\int \hat{a}^i(k)\ di}k)\ ,\ a.e.\ i \in [0,1]\} \tag{14}$$

That $k \in \mathcal{M}^\theta(k)$ thus means that there exists $(\hat{a}^i(k))$ which satisfies (12) and such that $k = \int \hat{a}^i(k)\ di$. Substitute this into (12) to see that $\hat{a}^i(k) \in S^{\theta,i}_{w(k),R(k)}(\hat{a}^i(k))$ which means that $\hat{a}^i(k)$ is an invariant distribution for household i. Comparing with Definition 1, we conclude that $k \in \mathcal{M}^\theta(k)$ if and only if k is an equilibrium.

Let $\mathcal{A}^i_k(K) \equiv \{\hat{a}^i \in \mathcal{P}(\bar{A}^i) : \hat{a}^i \in S^{\theta,i}_{w(k),R(k)}(\hat{a}_i \frac{k}{Q})\}$ where $\mathcal{P}(\bar{A}^i)$ is the set of probability measures on the compact range $\bar{A}^i \subseteq \mathbb{R}$ of the savings correspondence equipped with the weak $*$-topology, and $K > 0$. Since $S_{w,R,z^i}(a^i)$ is increasing and upper hemi-continuous in a^i, the adjoint Markov correspondence $S_{w,R}(\hat{a}^i)$ is type I and type II monotone and upper hemi-continuous in $\hat{a}^i$ (see the Appendix in Acemoglu and Jensen (2015)), so it follows from the fixed point comparative statics Theorem 3 in Acemoglu and Jensen (2015) that $\mathcal{A}^i_k(K)$ is type I and type II monotone in K^{-1}. By Theorem 4 in that same paper, $\int \mathcal{A}^i_k(\cdot)\ di$ has decreasing least and greatest selections. Since $\int \mathcal{A}^i_k(\cdot)\ di$ is convex valued by Richter's theorem (see Aumann (1965)), and a convex and real-valued correspondence whose least and greatest selections are decreasing must have a convex set of fixed points, $\mathcal{M}^\theta(k) = \{K : K \in \int \mathcal{A}^{\theta,i}_k(K)\ di\}$ is therefore convex. That the market correspondence $\mathcal{M}^\theta(k) = \{K : K \in \int \mathcal{A}^{\theta,i}_k(K)\ di\}$ is also upper hemi-continuous is seen by noting that its graph is $\{(k, K) : (K, k, K) \in \mathrm{Graph}[\int \mathcal{A}^{\theta,i}_k(K)\ di]\}$ where $\mathrm{Graph}[\int \mathcal{A}^{\theta,i}_k(K)\ di] = \{(K, k, Z) : Z \in \int \mathcal{A}^{\theta,i}_k(K)\ di\}$ is a closed set since $\int \mathcal{A}^{\theta,i}_k(K)\ di$ is upper hemi-continuous in k and K (this is shown by the same argument using now that $S_{w(k),R(k)}(\hat{a}^i \frac{k}{K})$ is upper hemi-continuous in $\hat{a}^i$ as explained above, as well as in k and K since $w(k)$ and $R(k)$ are continuous in k). That $\mathcal{M}^\theta(k)$ is compact follows now from boundedness (savings correspondences have compact ranges). Finally, $\mathcal{M}^\theta(k)$ begins above the 45 ° line and ends below it. The former is obvious since $f(0) = 0$ and therefore $\mathcal{M}^\theta(0) = \{0\}$. The latter is true since consumption is non-negative and therefore $\mathcal{M}^\theta(k) \leq Af(k)$, and the function on the right-hand-side eventually will lie below the 45 ° line (the production technology is effectively compact). ■

The market correspondence being convex-valued is a non-trivial property, in particular, it does *not* simply follow from a convexification argument as in Aumann (1965) (even though we are also implicitly using a convexication argument as part of the proofs). In fact, if S^i were not increasing (and consequently, the correspondence $\int \mathcal{A}^{\theta,i}_k(\cdot)\ di$ in the proof did not necessarily have decreasing least and greatest selections), the market correspondence would not necessarily be convex-valued.

4 Robust Comparative Statics in the Behavioral Neoclassical Growth Model

The previous section developed our general framework (the behavioral neoclassical growth model), and showed how this subsumes a broad range of non-standard preferences, biases, misperceptions, and near-rational behaviors. Lemma 1 then established that under Assumption 1, all of these household behaviors can be encoded in the market correspondence $\mathcal{M}^\theta$ of Definition

2 and that steady-state equilibria correspond to points where $\mathcal{M}^\theta$'s graph intersects with the 45° line. Our main result in this section will show that at the same level of generality, any change in the environment θ that is a local positive shock (defined below) leads to greater capital-labor ratio in the long run while a local negative shock leads to a lower capital-labor ratio.

Recall that the environment $\theta = (\theta^M, (P^i)_{i\in[0,1]})$ contains all of the exogenous variables, parameters and policy variables of the model as well as specifications of how beliefs about exogenous or endogenous objects are formed.[18] The *set of possible environments* is denoted by Θ and is taken to be an ordered set to facilitate our comparative static analysis. For any given environment $\theta \in \Theta$ the associated market correspondence $\mathcal{M}^\theta$ of Definition 2 then determines the set of (steady-state) equilibria $\{k^* : k^* \in \mathcal{M}^\theta(k^*)\}$ when Assumption 1 holds (Lemma 1). In the following definition, k^* is such a steady-state equilibrium given an (initial) environment $\theta^* \in \Theta$.

Definition 3 ***(Local Positive and Negative Shocks)*** *A change in environment from $\theta^* \in \Theta$ to $\theta^{**} \in \Theta$ is a local positive shock at k^* if there exists a $\tilde{k} \in \mathcal{M}^{\theta^{**}}(k^*)$ with $\tilde{k} \geq k^*$. If this inequality is reversed, the change in environment is a local negative shock at k^*.*

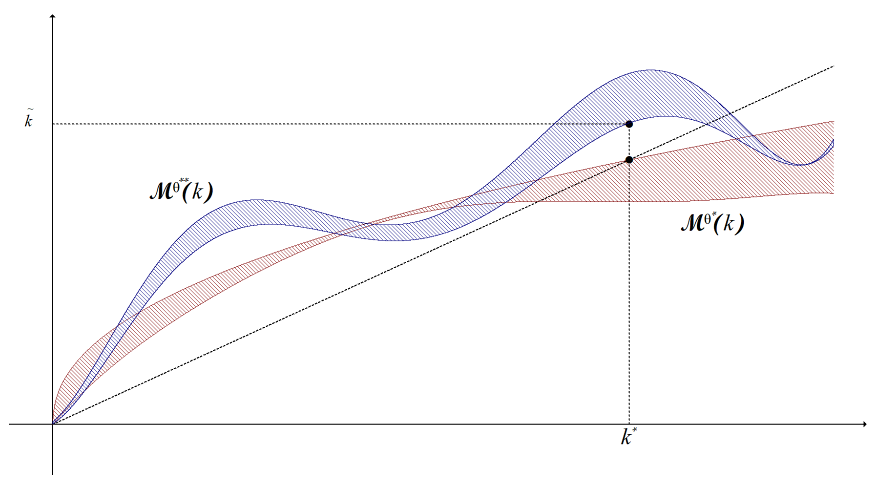

Figure 4: A local positive shock at k^*.

Intuitively, a local positive shock simply shifts the market correspondence up at the equilibrium k^* (Figure 4). Crucially, we show in Section 5.1 that this notion is equivalent to an increase in average (gross) savings at k^*.

[18]Here we are allowing savings correspondences to be set-valued, which is useful for two separate reasons. First, saving decisions are potentially non-unique. Second, even if saving decisions are uniquely determined, it is possible that multiple specifications of how beliefs are formed are consistent with steady-state behavior. In this latter case, the relevant set of limiting beliefs may even depend on the transition path of the economy (as in some learning models), but our focus will continue to be on the greatest and least equilibria.

Several things are important to note here. First, we have defined a local positive shock starting from the initial state state, k^*, for simplicity.[19] This is without loss of generality since throughout we focus on the consequences of a change in policy or parameters starting from an initial steady-state equilibrium.

Secondly, and more importantly, what makes this a local definition is that it is defined with reference only to the initial level of the capital-labor ratio (here k^*). In contrast, a "global positive shock" would impose a shift up of the market correspondents everywhere — that is, it would require $\mathcal{M}^{\theta^{**}}(k) > \mathcal{M}^{\theta^*}(k)$ for all k according to some set order (*e.g.*, the strong set order of Topkis (1978)). In contrast, as Figure 4 shows, our definition does not rule out the possibility that $\mathcal{M}^{\theta^{**}}(k) < \mathcal{M}^{\theta^*}(k)$ for $k \neq k^*$. In fact, even with this focus on a specific capital-labor ratio, k^*, ours is the weakest possible definition of a positive shock, because it does not even require that $\max \mathcal{M}^{\theta^{**}}(k^*) > \max \mathcal{M}^{\theta^*}(k^*)$, but only that there exists $\tilde{k} \in \mathcal{M}^{\theta^{**}}(k^*)$ with $\tilde{k} \geq k^* \in \mathcal{M}^{\theta^*}(k^*)$.[20]

Finally, in an economy with no uncertainty, a local positive (or negative) shock is simply about how households respond to a change in parameters or policy at the initial capital-labor ratio k^* or equivalently at a given vector of prices determined via (7)-(8) from k^* (with the caveat already noted in footnote 8). More generally, this is true whenever households' *belief formation*, *i.e.*, the rule that maps current observations and information about the model into beliefs about future variables, is fixed (*e.g.*, with rational expectations or with dogmatic expectations). When the belief formation varies over time (as we illustrate in the next footnote), a local positive shock is a statement about how households respond to a change in parameters or policy conditioning given that long-run belief formation or the equilibrium belief formation associated with θ^{**}.[21] This means, in particular, that when behavioral biases take the form of beliefs that change over time, what is relevant for our definition of a local positive shock is not initial beliefs, but beliefs after a sufficient amount of time has elapsed for the belief formation to be stationary.[22] That $\mathcal{M}^{\theta^{**}}$

[19] It is straightforward to give a definition of local positive (or negative) shock that applies starting from any capital-labor ratio. One that is equivalent to Definition 3 is the following: a change in environment from $\theta^* \in \Theta$ to $\theta^{**} \in \Theta$ is a local positive shock at k if there exists a $\tilde{k} \in M^{\theta^{**}}(k)$ with $\tilde{k} \geq \min\{k, \max M^{\theta^*}(k)\}$. If this inequality is reversed, the change in environment is a local negative shock at k. The equivalence follows immediately: since k^* is a steady state, $\min\{k^*, \max M^{\theta^*}(k^*)\} = k^*$. Intuitively, in this definition the first term of $\min\{k, \max \mathcal{M}^{\theta^*}(k)\}$ is included because what is relevant for a local positive shock is the part of the market correspondence that lies below the 45° line (before the change in parameter).

[20] For example, a definition that requires $\max \mathcal{M}^{\theta^*}(k) \geq \max \mathcal{M}^{\theta^{**}}(k)$ would imply Definition 3, but is clearly not implied by it. One drawback of a stronger definition would be that the equivalence between a local positive shock and an increase in average savings established in Section 5 would no longer be true, and as a consequence of this, the "if" part, but not the "only if", of Theorem 1 would apply with this stronger definition.

[21] Note that in the latter case, there will frequently be multiple admissible equilibrium belief formations which we cannot discriminate between given our focus on steady states. See footnote 18.

[22] For example, consider a representative household economy with the only deviation from the benchmark neoclassical model that the representative household has incorrect expectations in the short run, so when there is, say, a cut in the capital income tax rate, the representative household does not at first understand this (see, for example, Gabaix

should condition on how beliefs are formed in the long run can be seen from the following: if $\mathcal{M}^{\theta^{**}}$ were not conditioned on the equilibrium beliefs associated with θ^{**}, then its fixed points would not correspond to steady-state equilibria.

In the following two sections, we investigate various sufficient conditions that verify Definition 3.[23]

The next theorem provides our main result in the simplest case in which we focus on economies with a unique steady-state equilibrium. We provide generalizations of this result to settings with multiple steady states below.

Theorem 1 ***(Local Positive Shocks, Unique Steady State)*** *Assume that households satisfy Assumption 1. For environments $\theta^*, \theta^{**} \in \Theta$ let $k^* \in \mathcal{M}^{\theta^*}(k^*)$ and $k^{**} \in \mathcal{M}^{\theta^{**}}(k^{**})$ denote the associated (non-trivial steady state) equilibria and assume that these are unique. Then $k^{**} \geq k^*$ if and only if the change in environment from θ^* to θ^{**} is a local positive shock at k^*. Similarly, $k^{**} \leq k^*$ if and only if the change in environment from θ^* to θ^{**} is a local negative shock at k^*.*

Proof. Consider the case of a local positive shock.

Sufficiency: Since the change to θ^{**} is a local positive shock, there exists $\tilde{k} \in \mathcal{M}^{\theta^{**}}(k^*)$ with $\tilde{k} \geq k^*$. Since $\mathcal{M}^{\theta^{**}}$ ends below the diagonal, it must begin above and end below the 45° line on the interval $[k^*, +\infty)$. $\mathcal{M}^{\theta^{**}}$ is also upper hemi-continuous and convex valued (Lemma 1), hence it intersects the 45 degree line at some point k^{**} on $[k^*, +\infty)$. This yields a steady-state equilibrium $k^{**} \geq k^*$ given environment θ^{**}, and by assumption, this is the unique steady-state equilibrium.

Necessity: Assume that $k^{**} \geq k^*$ and that the change from θ^* to θ^{**} is not a positive shock at k^*. So $\sup \mathcal{M}^{\theta^{**}}(k^*) < k^*$ since the market correspondence is closed. But then since the market correspondence ends below the 45° line and is upper hemi-continuous and convex valued, $\mathcal{M}^{\theta^{**}}$ must intersect with the 45° at least twice on the interval $[k^*, +\infty)$. This contradicts that the economy has a unique interior steady state given θ^{**}.

The case of a local negative shock is proved by an analogous argument and is omitted. ■

(2017), Section 5). In this case, the initial belief formation differs from the way beliefs are formed after some amount of time, given by some large enough T, has elapsed because the household initially misperceives the change in policy. In fact, it may well be that the change in policy corresponds to a local negative shock if we condition on initial beliefs but to a local positive shock when we condition on long run beliefs. Then the relevant definition of the local shock is the latter. Thus after the policy shock, we will first experience a negative impact effect on average savings. Nevertheless, the change in policy is still a local positive shock and the long-run capital-labor ratio will increase.

[23]It is useful to bear in mind that in one instance, namely the case with a representative household, we already know the answer from Section 2: The change in environment is a local positive shock if and only if it raises the representative household's savings given k^*. See equation (2) and the surrounding discussion of its relationship with local positive shocks.

Theorem 1 provides our sharpest result focusing on the case where the steady-state equilibrium is unique before and after the environment changes. It shows how all we need to know is that the change in environment is a local positive shock. Given the relatively weak regularity conditions we have imposed on the market correspondence, we can then conclude that the full equilibrium effect will be to increase the capital-labor ratio in the new steady state. Conversely, for a local negative shock, the new steady state will always involve a lower capital-labor ratio.

The intuition for this result was already provided in the Introduction. Briefly, even though there is a large amount of heterogeneity and potential biases and mistakes, the aggregate behavior cannot go the wrong way, because of the "market discipline" coming from the fact that some agents must be increasing their savings as a result of the initial impetus coming from the local positive shock, and this initial effect cannot be undone by the indirect equilibrium responses. It is also

useful to spell out why these indirect effects can never win out. Note first that, by definition, a local positive shock increases savings at the initial capital-labor ratio. Therefore, the only way we may end up with a paradoxical result where the new steady-state equilibrium involves a lower capital-labor ratio than the initial one is when this higher level of savings induces so much dissaving from some households that the indirect equilibrium response more than offsets the initial impetus. (Such a paradoxical result is impossible if the equilibrium responses led to more saving than dissaving). But if this indirect effect did indeed overwhelm the initial local positive shock, that would mean that in the new steady state there would be a lower capital-labor ratio and thus there would be no reason for the dissaving to offset the initial impetus. This contradicts the possibility that the indirect effect could more than offset the initial impact coming from the local positive shock.

The rest of this section is devoted to generalizing this result to cases in which we do not necessarily have uniqueness. Note that as an immediate consequence of Lemma 1, the set of equilibria is always non-empty and compact.[24] So even if the non-trivial equilibrium is not unique, the set of equilibria is nonetheless guaranteed to reside in a closed interval which allows us to study the interval marked by the greatest and least equilibria.

The next theorem shows that the conclusions of Theorem 1 directly carry over to the case in which the shocks we are considering are "small" (meaning that we can choose them to be small enough given the setting).[25]

[24]The set of steady states is given by the intersection between a compact set (the graph of the market correspondence) and a closed set (the 45° line). The intersection is therefore compact and also non-empty (since the graph of the market correspondence is connected and begins above and ends below the 45° line).

[25]Throughout by the least steady state we are referring to the least non-trivial steady state, thus excluding $k = 0$.

Theorem 2 ***(Greatest and Least Steady States under Multiplicity I)*** *Let the assumptions of Theorem 1 hold and define* $k^*_- = \inf\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ *as the least steady state and* $k^*_+ = \sup\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ *as the greatest steady state when the environment is* $\theta^* \in \Theta$*, and analogously* k^{**}_- *and* k^{**}_+ *when the environment is* $\theta^{**} \in \Theta$*. Assume in addition that* $\mathcal{M}^\theta$ *is upper hemi-continuous in* $\theta \in \Theta$ *(where now* Θ *is a topological space). Consider an infinitesimal change in the environment to* θ^{**}*. Then,* $k^{**}_- \geq k^*_-$ *if and only if the change in environment is a local positive shock at* k^*_-*, and* $k^{**}_+ \geq k^*_+$ *if and only if the changing environment is a local positive shock at* k^*_+*.*

Proof. Since the market correspondence is compact-valued, a sufficiently small change in the environment can lead to existing equilibria disappearing but not to the creation of new equilibria. In particular, no new equilibrium can be created below the least equilibrium which must therefore increase by the argument used to prove Theorem 1. This argument obviously also applies to the the greatest equilibrium; and in both cases necessity follows by the argument from Theorem 1 as well. ■

If there are multiple equilibria and the change in environment is not small in the sense of the previous result (or we are unwilling or unable to place a topology on the set of possible environments Θ), we can still identify how the greatest equilibrium will respond when the change in environment is a local positive shock and how the least equilibrium will respond when the change in environment is a local negative shock.

Theorem 3 ***(Greatest and Least Steady State under Multiplicity II)*** *Let the assumptions of Theorem 1 hold and consider* $k^* = \sup\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ *(the greatest steady state) of the environment* $\theta^* \in \Theta$*. Then if a change from* θ^* *to a new environment* $\theta^{**} \in \Theta$ *is a local positive shock at* k^**, the economy's greatest steady state increases, i.e.,* $\sup\{k : k \in \mathcal{M}^{\theta^{**}}(k)\} \geq k^*$*. Analogously, consider* $k^* = \inf\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ *(the least steady state) of the environment* $\theta^* \in \Theta$*. Then if a change from* θ^* *to a new environment* $\theta^{**} \in \Theta$ *is a local negative shock at* k^**, the economy's least steady state decreases, i.e.,* $\inf\{k : k \in \mathcal{M}^{\theta^{**}}(k)\} \leq k^*$*.*

Proof. Let k^* denote the greatest steady state. Repeating the argument used to prove the "sufficiency" part of Theorem 1, $\mathcal{M}^{\theta^{**}}$ must have a fixed point on $[k^*, +\infty)$. The result for the least steady-state is proved analogously. ■

Finally, we can pin down the behavior of both the greatest and least steady states if we are willing to impose that the change in environment is a local positive shock at the greatest steady state and that the reverse change is a local negative shock at the least steady state (and analogously for a local negative shock).

Theorem 4 ***(Greatest and Least Steady States under Multiplicity III)*** *Let the assumptions of Theorem 1 hold and denote by $k_-^\theta = \inf\{k : k \in \mathcal{M}^\theta(k)\}$ the least steady state, and by $k_+^\theta = \sup\{k : k \in \mathcal{M}^\theta(k)\}$ the greatest steady state for an environment $\theta \in \Theta$. Consider two environments $\theta^*, \theta^{**} \in \Theta$. Then if the change from θ^* to θ^{**} is a local positive shock at $k_+^{\theta^*}$, and the change from θ^{**} to θ^* is a local negative shock at $k_-^{\theta^{**}}$, the economy's greatest and least steady states increase as the environment changes from θ^* to θ^{**}.*

Proof. Since the change from θ^{**} to θ^* is a local negative shock at $k_-^{\theta^{**}}$, we may apply Theorem 3 with θ^* and θ^{**} interchanged to conclude that $k_-^{\theta^*} \leq k_-^{\theta^{**}}$. That the greatest steady state must increase follows directly from Theorem 3. ■

Appendix A contains additional results along the lines of the previous theorems. Although important for theoretical applications, the details are less central to our substantive results, hence its relegation to Appendix A. In addition, we also provide there a detailed comparison with related equilibrium comparative statics results in Milgrom and Roberts (1994) and Acemoglu and Jensen (2013).

5 Local Positive Shocks and Individual Behavior

Recall from Section 2 (*e.g.*, equation (2) and the surrounding discussion) that in a deterministic environment with a representative household the market correspondence is given by the representative household's gross savings less government consumption and waste from distortions. This implies that a change in the environment from θ^* to θ^{**} is a local positive shock at the equilibrium $k^* \in \mathcal{M}^{\theta^*}(k^*)$ if and only if the direct effect on savings is positive, that is,

$$s^{\theta^{**}}_{w^{\theta^{**}}(k^*),R^{\theta^{**}}(k^*)}(k^*) \geq s^{\theta^*}_{w^{\theta^*}(k^*),R^{\theta^*}(k^*)}(k^*) \tag{15}$$

where we have conditioned factor prices on the environment for emphasis, though in what follows we will suppress this dependence for notational convenience whenever doing so can cause no confusion.

The purpose of this section is to extend the previous equivalence between changes in savings and local positive shocks to the general setting of this paper. This section also contains two key examples that formally demonstrate two of the main conclusions that were highlighted in the Introduction (Sections 5.3-5.4). To simplify the exposition in this section, we assume throughout that the function describing non-rebated tax income and waste, $G(k)$, is given.

5.1 Necessary and Sufficient Conditions in Terms of Savings

We now derive an analogue of equation (15) describing a local positive shock in the general environment with uncertainty and heterogeneous agents. To do so, we need a bit of additional terminology and notation. Consider an equilibrium $k^* \in \mathcal{M}^{\theta^*}(k^*)$, where as in the previous section $\theta^* \in \Theta$ should be thought of as the default environment. Equilibrium prices $w(k^*)$ and $R(k^*)$ continue to be given by (7)-(8). Let us fix these market prices (*i.e.*, fix k^*) but change the environment to θ^{**}, and let $\mathbb{E}[\hat{a}^{\theta^{**},i}(k^*)]$ denote the mean asset holdings of household i in the new environment (but starting at the capital-labor ratio k^* and the associated prices).[26] Intuitively, $\mathbb{E}[\hat{a}^{\theta^{**},i}(k^*)]$ is the answer we get from asking the agent how much she expects to save on average if the environment changes to θ^{**} from the default environment θ^* when all prices (the capital-labor ratio) remain forever the same. To simplify notation, we assuming this subsection that this mean asset holding is uniquely determined for all households, but as explained in Remark 1, the results here can be easily extended to the case with multiple equilibrium asset holdings.

Definition 4 ***(Changes in Mean Asset Holdings)*** *Let $k^* \in \mathcal{M}^{\theta^*}(k^*)$ be a steady in the environment $\theta^* \in \Theta$. The population's mean asset holdings increase at k^* when the environment changes from θ^* to θ^{**} if $\int \mathbb{E}[\hat{a}^{\theta^{**},i}(k^*)]\, di \geq \int \mathbb{E}[\hat{a}^{\theta^{*},i}(k^*)]\, di$. If the inequality is reversed, the population's mean asset holding decreases at k^* when the environment changes from θ^* to θ^{**}.*

The definition is intuitive: We simply average over the asset holdings (or gross savings) of households in the old and new environments at given after-tax prices and trace the direction of change. The next proposition connects this with local positive shocks showing that, in equilibrium, the two are equivalent. While this result is intuitive in light of the overview in Section 2, it is far from trivial because the market correspondence is formally defined by solving a fixed point problem and equalizing $\mathcal{M}^{\theta}(k)$ with the means of the set of solutions (see (14) in the proof of Lemma 1).

Proposition 1 ***(Mean Asset Holdings and Local Shocks)*** *Assume that households satisfy Assumption 1, and let $k^* \in \mathcal{M}^{\theta^*}(k^*)$ be either the least steady state $\inf\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ or the greatest steady state $\sup\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ given some $\theta^* \in \Theta$. Then the change in environment from θ^* to θ^{**} is a local positive shock at k^* if and only if the population's mean asset holdings increase at k^* when the environment changes from θ^* to θ^{**}. Similarly, the change in environment is a local negative shock if and only if the population's mean asset holdings decrease at k^*.*

[26]Note that $\hat{a}^{\theta^{**},i}(k^*)$ is determined by $\hat{a}^{\theta^{**},i}(k^*) \in S^{\theta^{**},i}_{w^{\theta^*}(k^*),R^{\theta^*}(k^*)}(\hat{a}^{\theta^{**},i}(k^*))$.

Proof. From the proof of Lemma 1 we have $\mathcal{M}^{\theta}(k) = \{K : K \in F_k(K, \theta)\}$ where

$$F_k(K,\theta) = \{\int a^i \, di : a^i \in S^{\theta,i}_{w(k),R(k)}(\frac{a^i}{K}k) \text{ a.e. } i\}$$

Note that $F_{k^*}(\cdot, \theta^*)$ as well as $F_{k^*}(\cdot, \theta^{**})$ are upper hemi-continuous, convex valued, and begin above and end below the diagonal (the latter follows from the fact that it is decreasing in K, see the proof of Proposition 1). Let k^* be the the greatest equilibrium. By the same argument as the one used to prove sufficiency in Theorem 3, it follows that if there is a $\hat{k} \in F_{k^*}(k^*, \theta^{**})$ with $\hat{k} \geq k^*$, then there exists $K \in F_{k^*}(K, \theta^{**})$ with $K \geq k^*$. But since $F_{k^*}(k^*, \theta^{**})$ is the population's mean asset holdings at k^* in environment θ^{**}, and we have assumed this is greater than or equal to k^*, we have $K \in \mathcal{M}^{\theta^{**}}(k^*) \Leftrightarrow K \in F_{k^*}(K, \theta^{**})$. So the change in environment is a local positive shock. This argument also applies if k^* is the least equilibrium since F is decreasing in K. We remark that in the multiple steady-state asset distributions case discussed in Remark 1 below, $F_{k^*}(k^*, \theta^{**})$ is not single-valued but once again the argument goes trough as long as F's maximum is greater than or equal to k^*. To see that an increase in mean asset holdings is also necessary for a local positive shock use that if there does not exist $\hat{k} \in F_{k^*}(k^*, \theta^{**})$ with $\hat{k} \geq k^*$, then because $F_k(K, \theta^{**})$ is convex valued with least and greatest selections that are decreasing in K, there is not a $K \in F_{k^*}(K, \theta^{**})$ with $K \geq k^*$, and so the change from θ^* to θ^{**} is not a local positive shock at k^*. ■

Remark 1 ***(Multiplicity of Equilibrium Asset Distributions)*** *If agent i's steady-state asset distribution is not uniquely determined from k, we consider the greatest mean asset holdings: $A^{\theta,i}_{+}(k) = \sup\{\mathbb{E}[\hat{a}^i] : \hat{a}^i \in S^{\theta,i}_{w(k),R(k)}(\hat{a}^i)\}$. From here we define the greatest average asset holdings across the agents (given θ and the steady state k): $A^{\theta}_{+}(k) = \int A^{\theta,i}_{+}(k) \, di$. With these in hand, the following natural generalization of Proposition 1 holds: Let $k^* \in \mathcal{M}^{\theta^*}(k^*)$ be either the least steady state* $\inf\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ *or greatest steady state* $\sup\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ *given some $\theta^* \in \Theta$. Then the change in environment from θ^* to θ^{**} is a local positive shock at k^* if and only if $k^* \leq A^{\theta^{**}}_{+}(k^*)$ (see the proof of Proposition 1). Note that trivially the left-hand side of this inequality, k^*, is the average asset holding across the households at the steady state k^*. So, the necessary and sufficient condition is that the greatest average asset holding after the change in environment is above the average asset holdings before the change.*

5.2 The Case of Uniform Direct Effects

A particularly simple case of changes in the environment that are covered by Definition 4 is when almost every individual's direct effects go in the same direction — which is clearly sufficient and very far from being necessary for a local positive or negative shock.[27]

Definition 5 ***(Individual Direct Effects)*** *Let k^* be an equilibrium given θ^* and denote by $\hat{a}^i$ household i's associated steady state assets. We say that household i's asset holdings increases (or that its direct effect is positive) at k^* when the environment changes from θ^* to θ^{**} if*

$$S^{\theta^{**},i}_{w^{\theta^{**}}(k^*),R^{\theta^{**}}(k^*),z^i}(a^i) \geq S^{\theta^*,i}_{w^{\theta^*}(k^*),R^{\theta^*}(k^*),z^i}(a^i) \text{ a.e. } z^i \in Z^i \text{ and } a^i \in \mathrm{Support}(\hat{a}^i). \tag{16}$$

If the inequality is reversed, we say instead that the household's asset holdings decreases when the environment changes from θ^* to θ^{**}.

In Acemoglu and Jensen (2015) we imposed such uniform direct effects and also required that individuals' savings levels increase for all k and so for all possible prices (rather than just the initial capital-labor ratio as we are doing here and in Definition 4). Appendix A provides additional discussion of the relationship of our approach here to our and others' previous work.

Note also that in the deterministic case, $\mathrm{Support}(\hat{a}^i)$ in the definition contains just a single element (namely the economy's steady state). So Definition 5 reduces to (15) when there is no uncertainty and the savings correspondence is single-valued. As an example, consider individuals with idiosyncratic labor endowment shocks who may be borrowing constrained as in Aiyagari (1994). Light (2017) shows that such agents will increase savings (and asset holdings) if the interest rate increases in the standard case with CRRA preferences and rate of risk aversion weakly below unity (Light (2017), Theorem 1). Hence individual direct effects are positive. See also Acemoglu and Jensen (2015) who identify a variety of changes in environments whose direct effects are positive (in the strong sense of holding for all prices as mentioned a moment ago).

Proposition 2 ***(Uniform Increases in Asset Holdings are Local Positive Shocks)*** *Let $k^* \in \mathcal{M}^{\theta^*}(k^*)$ be either the least equilibrium $\inf\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ or the greatest equilibrium $\sup\{k : k \in \mathcal{M}^{\theta^*}(k)\}$ given some $\theta^* \in \Theta$. If almost every household's asset holdings increases at k^* when the environment changes from θ^* to $\theta^{**} \in \Theta$, then the change in environment from θ^* to θ^{**} is a local positive shock at k^*. Similarly, if almost every household's asset holdings decreases at k^* when the environment changes from θ^* to $\theta^{**} \in \Theta$, then the change in environment from θ^* to θ^{**} is a local negative shock at k^*.*

[27]More explicitly, the statement in (16) is that $\sup S^{\theta^{**},i}_{w(k^*),R(k^*),z^i}(a^i) \geq \sup S^{\theta^*,i}_{w(k^*),R(k^*),z^i}(a^i)$ and $\inf S^{\theta^{**},i}_{w(k^*),R(k^*),z^i}(a^i) \geq \inf S^{\theta^*,i}_{w(k^*),R(k^*),z^i}(a^i)$ (this is the standard ordering of sets which was also used in Section 3). Note also that we suppress the dependence of the wage and interest on the environment as mentioned immediately after (15) above.

Proof. Since savings correspondences are increasing in assets under Assumption 1, it follows directly from Theorems 3-4 in Acemoglu and Jensen (2015) that (almost) every household's mean asset holdings must increase. The conclusion then follows from Proposition 1. ■

The sufficient conditions provided in this proposition for local positive (or negative) shock are not the only possible ones. One can alternatively use stochastic dominance relations (see Jensen (2018)) to establish that a change in environment is a local positive shock without imposing uniformity of direct effects.

5.3 Indeterminacy of Individual Behavior

This section illustrates that although, as our main results show, comparative statics of some generality can be established for aggregate outcomes (such as the capital-labor ratio and consequently prices), little can be said in general about how individuals will respond to changes in the environment, including to policy changes. Even in the very special case where *every* household's mean asset holdings increase at the initial capital-labor ratio k^*, *i.e.*, where $\mathbb{E}[\hat{a}^{\theta^{**},i}(k^*)] \geq \mathbb{E}[\hat{a}^{\theta^*,i}(k^*)]$ for all $i \in [0,1]$, a strictly positive measure subset $A \subset [0,1]$ of households may end up reducing their gross savings and mean asset holdings in the new equilibrium k^{**}, that is, $\mathbb{E}[\hat{a}^{j,\theta^{**}}(k^{**})] < \mathbb{E}[\hat{a}^{j,\theta^*}(k^*)]$ for $j \in A$. In the benchmark model with forward-looking rational households this may be the case because of countervailing income and substitution effects. In this subsection we demonstrate that when we depart from this benchmark by allowing richer models of consumption and saving behavior, even less can be said about individual behavior.

We now illustrate this claim in the context of the simple example of a reduction in capital income taxes as in our discussion in Section 2. Recall from our discussion there that $\Delta(k)$ is defined as the change in the rate of return to capital due to the change in tax policy from $\hat{\tau}(k)$ to $\tau(k)$. Here we allow this change in the rate of return to capital to be individual specific, and write it as $\Delta^j(k) = (\tau^j(k) - \hat{\tau}^j(k))f'(k)$. This might be because taxes vary across households. We also assume for concreteness that there will be a corresponding decline in government consumption as taxes on capital income decline. Finally, let us assume that the (non-trivial) equilibrium is unique and all saving correspondences are single-valued (all of the simplifying assumptions stack the deck against establishing that little can be said about individual behavior). Observe next that if a household j reduces savings at *every* state of the world z^j and asset level a^j, i.e.,

$$s^{j,\theta^{**}}_{w(k^*),R(k^*)+\Delta^j(k^*),z^j}(a^j) < s^{j,\theta^*}_{w(k^*),R(k^*),z^j}(a^j) \text{ for all } z^j \text{ and } a^j, \tag{17}$$

then it must reduce its mean asset holdings in equilibrium.[28]

Let us now start with the benchmark case where all households have rational expectations and maximize geometrically (dynamically consistent) discounted CRRA objectives with rates of risk aversion below unity. If $\Delta^j(k^*) \geq 0$ for all j, direct effects will then be uniformly positive (see the paragraph after Definition 5), so from Proposition 2 this is a local positive shock. Therefore, from Theorem 1, we have $k^{**} > k^*$, and the equilibrium wage rate will increase and the equilibrium interest rate will decline, that is, $w(k^{**}) > w(k^*)$ and $R(k^{**}) < R(k^*)$.[29] It is useful to note that any household that does not benefit from the tax reduction (i.e., for which $\Delta^j(k^*) = 0$) will display zero direct effect. But critically, such a household will still adjust its savings as a result of the indirect (equilibrium) effects originating from the changes in the equilibrium wage and interest rate just described. Because in this benchmark case household savings are increasing in the interest rate (holding assets constant), we have $s^j_{w(k^*),R(k^{**}),z^j}(a^j) < s^j_{w(k^*),R(k^*),z^j}(a^j)$ and the lower equilibrium interest rate will push towards lower gross savings for these directly unaffected households. Similarly, provided that the household in question is not borrowing constrained, the response to the increase in the equilibrium wage is negative too: $s^j_{w(k^{**}),R(k^{**}),z^j}(a^j) < s^j_{w(k^*),R(k^{**}),z^j}(a^j)$.[30] Combining these two inequalities, we can conclude that (17) must hold. This discussion thus establishes that in the benchmark case with forward-looking rational households (and dynamically consistent objectives), any household j that is not borrowing constrained and is not affected by the tax reduction must reduce its mean asset holdings in equilibrium — even as gross savings averaged across households necessarily increases since $k^{**} > k^*$ from Theorem 1.

A similar construction shows that behavioral factors make predictions about individual responses even more challenging. Suppose that the capital income tax is reduced for all households and income and substitution effects are such that when all households fully optimize they will all increase gross savings (as in the case of uniform direct effects studied in the previous sub-

[28]This can be proved by the exact same argument as the one used to prove Proposition 2.

[29]Note that we have here put strict inequalities. Due to space limitations, we have throughout the paper avoided making a clear distinction between "weak" changes and "strict" changes in equilibrium. But by simply looking at a market correspondence (*e.g.* Figure 3), it is clear that the equilibrium change will in fact be strict unless the maximal element of the market correspondence remains exactly the same at k^* as the environment changes. This cannot happen if any positive measure of agents strictly increase their mean asset levels with the change in environment (keeping everything else fixed as described in Section 5.1). In the current example, this is guaranteed for any subset of households whose rate of risk aversion is strictly below unity and who also experiences a strict reduction in the capital income tax.

[30]Under forward looking behavior and with no borrowing constraint, a permanent increase in w is equivalent to increasing wealth by $\frac{wz^j}{R}$ and adjusting consumption correspondingly: $s^j_{w,R,z^j}(a^j) = s^j_{0,R,0}(a + \frac{wz^j}{R}) - \frac{wz^j}{R}$. Clearly savings must therefore be decreasing in w (unless the marginal propensity to save is above 1 which is a case we can safely discard in equilibrium). See also footnote 8 in Cao and Werning (2017) who use this same observation to lump labor income into wealth (*i.e.*, work with the savings function $s^j_{0,R,0}(a + \frac{wz^j}{R})$ in place of $s^j_{w,R,z^j}(a^j)$).

section). But suppose, instead, that a subset of households have systematically incorrect beliefs and misperceive the tax cut as unchanged capital income taxes. Then under these mistaken beliefs, the same argument as in the previous paragraph applies and shows that the gross savings of this set of agents will decline rather than increase. This simple example thus illustrates that optimization mistakes will make predicting individual behavior even more difficult. Though the case of systematically incorrect beliefs about the tax cut is extreme, it is straightforward to introduce other constraints on optimization or behavioral biases which will deliver the same point — even in cases where we would have been able to characterize individual behavior with fully optimizing (etc) agents, richer behavioral preferences add another layer of indeterminacy.

This indeterminacy of individual behavior further underscores the power of our approach: a strategy attempting to determine how aggregates change based on individual changes would not be able to make progress for the simple reason that nothing much can be said about individual behavior. Our results show that even though individual behavior is indeterminate, we can still in considerable generality say how aggregates change in response to local positive or local negative shocks.

5.4 When Behavioral Biases Matter

The results presented so far do not imply that behavioral biases are unimportant. Rather, they establish that in the context of the one-sector neoclassical growth model, if despite these biases a change in policies or parameters of the model is a local positive shock, then the long-run impact on the capital-labor ratio will be positive. Therefore, behavioral biases do not matter for the direction of long-run comparative statics *provided that* they do not change whether an initial impetus is a local positive or negative shock. But conversely, in this subsection we show that, with a very similar reasoning, behavioral biases matter greatly — and change the direction of comparative statics — when they alter whether a change in policies or parameters is a local positive or negative shock.

This is straightforward to see using a slight modification of the example from the previous subsection. Suppose again that there is a cut in the capital income tax rate, now for all households, but differently from before, suppose also that the proceeds of taxes are being rebated to households in a lump sum fashion, so the tax cut is accompanied with a reduction in transfers. If all households have rational expectations and dynamically consistent recursive preferences, this policy change would be a local positive shock and thus increase the long-run capital-labor ratio. Suppose, instead, that a fraction $\alpha \in (0, 1)$ have systematic (permanent) misperceptions and do not understand the implications of the capital income tax cut for their after tax returns (even after an arbitrary number of periods), but do perceive the reduction in their non-capital income resulting from lower transfers; they may consequently reduce their savings. As a result, if this behavioral group of consumers have sufficiently high marginal propensity to save out of transfers, the cut in capital income tax may reduce their gross savings and turn the reduction in capital income tax into a local negative shock rather than the local positive shock that it would have been absent these behavioral consumers. But then by Theorem 1, the long-run capital labor ratio will decline — rather than increase — in response to this policy change.

Though this simple example may appear too simplistic or too extreme, its message is much more general. In the next proposition, we provide a more general result in this direction, showing that for any fraction $\alpha > 0$ of behavioral agents, a cut in the capital income tax rate can become a local negative shock even though it is a local positive shock absent the behavioral agents. For concreteness, we suppose that the non-behavioral agents with rational expectations have general recursive preferences as in Epstein and Hynes (1983) or Lucas and Stokey (1984), and that the behavioral agents perceive the tax reduction as an indication that "better economic times lie ahead", which makes them reduce precautionary savings so that $S^{\theta^{**}}_{w,\tilde{R}} < S^{\theta^{*}}_{w,\tilde{R}}$ for all w and $\tilde{R}$.

Proposition 3 *Suppose that a fraction $\alpha \in (0,1)$ households have behavioral preferences and reduce their savings in response to a cut in the capital income tax. The remaining households have rational expectations and recursive, dynamically consistent preferences as in Epstein and Hynes (1983) or Lucas and Stokey (1984). Then for any $\alpha > 0$, there exist recursive preferences for the remaining $1 - \alpha \in (0,1)$ households such that despite their increased savings, the tax cut is a local negative shock and thus the steady-state capital labor ratio decreases from k^* to $k^{**} < k^*$ following the tax reduction.*

Proof. See Appendix B. ■

6 Foundations of Individual Behavior

With the exception of Sections 5.2-5.4, we have so far taken consumption and savings correspondences as primitives in order to nest a wide range of behavioral models. While this strategy has the advantage of simplicity (we could directly impose Assumption 1), it begs the question whether this assumption is likely to be verified in interesting applications, including in the behavioral models we discussed in Section 3.1. This is the question we turn to in this section. Our conclusion is that with all of the behavioral preferences discussed or mentioned so far, Assumption 1 is satisfied, and therefore, these models are indeed naturally covered by our results.

We start with a brief summary before presenting our formal analysis. The key Definition 6 provides our main microfoundation of savings functions (a savings correspondence is simply defined as the union of all savings functions). The set of behavioral models covered by this definition is extensive and includes, among others, models incorporating various types of uncertainties and general beliefs (which may feature ambiguity in the sense of Gilboa (1987) and Schmeidler (1989)). The main restriction imposed in this definition is that behavior must be "time-stationary", meaning that it does not depend on calendar time (given the relevant state variables). With any dynamically consistent objective, this follows automatically (Strotz (1956)), and the set of savings functions determined in Definition 6 will therefore coincide with the savings functions implied by standard recursive dynamic programming formulations. In this category we find non-additive objectives as in Epstein and Zin (1989), and as a special case therefore Kreps and Porteus (1978) (see Example 6). Extensions of these models to ambiguity are also covered. In the deterministic case the definition includes recursive utilities a la Koopmans (1960) (see also Epstein and Hynes (1983) and Lucas and Stokey (1984)). But crucially, as we explain below in Example 4, behavioral models featuring mistakes or approximate rationality (including the Luce (1959)-model and satisficing behavior a la Simon (1956)) fit into this dynamically consistent category as well.[31]

If the objective is dynamically inconsistent, time-stationarity, and thus Definition 6, still apply, but now the relevant saving functions are given by the (Bayesian) Nash equilibria of the game played by households' temporal selves (*e.g.*, Phelps and Pollak (1968), Laibson (1996, 1997), Balbus, Reffett and Wozny (2015)). The obvious example is models of delay discounting (*e.g.*, hyperbolic or quasi-hyperbolic discounting as in Example 1), but dynamic inconsistency arises naturally in a number of other behavioral consumption decision models as well for reasons that are otherwise unrelated to discounting. For example, with "incorrect" beliefs/expectations (Example 3), a self at any date will generally observe a different outcome than the one that previous selves' foresaw and based their decisions on (in this case, dynamic inconsistency is embedded in the belief structure). As a second example, if agents' time horizon is of length $T < \infty$ — because of myopia or as a result of sparsity constraints as in (Gabaix (2014, 2017)) — selves at different dates will not "agree" on an overall objective which again leads to dynamic inconsistency (Example 2).

In the dynamically consistent case we show in Proposition 5 that savings correspondences satisfy Assumption 1 under the most general non-additive specification of Epstein and Zin (1989) if consumption at different dates are Edgeworth-Pareto complements in the sense of Chipman (1977), *i.e.*, if the marginal utility of future consumption is non-decreasing in current consumption. In the dynamically inconsistent case, a little more care is necessary. Nonetheless, Proposi-

[31] It is possible to "pair" approximate rationality with dynamic inconsistency, too, as in Example 4.

tion 4 establishes that if the underlying utility function exhibits weak separability in current and future consumption (not necessarily over the entire consumption stream), then once again savings correspondences will satisfy Assumption 1. We are not familiar with any behavioral model that simultaneously features dynamical inconsistency and non-separability between consumption of the current self and the consumption of future selves, but for such a model a different line of argument will need to be developed.

We end this section by showing that at the same level of generality as the previous results, it is possible to establish that some changes in preferences are positive shocks. In particular, we show that when a household becomes more patient, it increases its savings, and thus a shift in preferences towards greater patience is a positive shock for all the preferences we are considering here.

6.1 Time-Stationary Saving Correspondences

At any date t, household i is influenced by z_t^i, where z_t^i is a Markov process with invariant distribution μ_{z^i} (see Section 3.3). For concreteness, we assume here that $z^i = (\epsilon^i, l^i, \sigma^i)$ where l^i is the household's labor endowment, ϵ^i is a random utility parameter and σ^i an index of the extent of uncertainty faced or perceived by agent i. In addition, z^i could also contain random shocks to capital income, to taxes, etc. We also simplify the exposition by assuming that there are no direct (lump sum) taxes/transfers, so given assets a_t^i at date t, household i's income equals $(1 + R_t)a_t^i + w_t l_t^i$ where w_t and R_t are the after-tax wage and interest rates determined from (7) and (8).[32] Given its income at date t, the household chooses (gross) savings a' and consumes $c_t = (1 + R_t)a^i + w_t l_t^i - a'$. At date $t + 1$, the process repeats itself given the new asset level $a_{t+1}^i = a'$, and R_{t+1}, w_{t+1}, and the next realization of z^i, z_{t+1}^i.

At every date, the household's objective is to maximize utility conditioned on its beliefs (or expectations) about future variables $(R_\tau, w_\tau, z_\tau^i)_{\tau=t+1}^\infty$ and its own anticipated future savings behavior. Beginning with the former, recall from Section 3.3 that θ^M denotes a complete description of the model, including all future taxes and the stochastic process governing z_t^i. Beliefs are formed by combining this information with current observations through a mapping $P^i : (w, R, z^i, \theta^M) \mapsto P^i(\cdot; w, R, z^i, \theta^M)$ where $P^i(\cdot; w, R, z^i, \theta^M)$ is a probability measure, and we simplify the notation by writing it as $P^i(\cdot)$. For any measurable set B, the household thus believes that with probability $P^i(B) \in [0, 1]$ the future sequence $(w_\tau, R_\tau, z_\tau^i)_{\tau=t+1}^\infty$ lies in B (see the paragraph prior to Condition C below for details on the measurable space this statement refers to). In the notation of the previous sections, the "environment" is thus $\theta = (\theta^M, (P^i)_{i\in[0,1]})$. Note

[32]With lump sum income taxes/transfer $T_t^i \in \mathbb{R}$ at date t, income will instead be $(1 + R_t)a_t^i + w_t l_t^i + T_t^i$. The beliefs described next will then implicitly include beliefs about these lump sum taxes/transfers.

that since households form beliefs about future after-tax prices as well and these depend on taxes and distortions as described in Section 3.2, P^i implicitly incorporates beliefs about future taxes (tax policy) and distortions. Crucially, we do not impose rational expectations or that household beliefs are correct conditioned on θ^M; in fact, beliefs could be completely independent of the model θ^M as in the case of adaptive expectations, and expectations may temporarily or even permanently deviate from actual outcomes (the latter would be an instance of systematic misperceptions, see Section 3.1 and Example 3 below).[33] Finally note that we are not assuming that beliefs are additive measures, in particular, households may entertain multiple simultaneous beliefs about the future (Gilboa (1987), Schmeidler (1989), Gilboa and Schmeidler (1995)).

Apart from beliefs, the date t savings decision of a household's "current self" also depend on the anticipated savings behavior of its "future selves" (though as will become clear, the random utility parameter implies that the current self may be uncertain — or even wrong — about future selves' behavior). Thus savings $a^{i,\prime}$ at date $t = 0$ implies an anticipated random consumption stream that depends not only on beliefs but also on anticipated future behavior: $c_0 = (1 + R_0)a^i + w_0 l_0^i - a^{i,\prime}$, $c_1 = (1 + R_1)a^{i,\prime} + w_1 l_1^i - s_{w_1,R_1,z_1^i}(a')$, $c_2 = (1 + R_2)s_{w_1,R_1,z_1^i}(a^{i,\prime}) + w_2 l_2^i - s_{w_2,R_2,z_2^i}(s_{w_1,R_1,z_1^i}(a^{i,\prime}))$, $\ldots$; where $s_{w,R,z^i}(\hat{a}^i)$ is the savings function of a "future self" with assets $\hat{a}^i$ and beliefs $\tilde{P}^i(\cdot) = P^i(\cdot; w, R, z^i, \theta^M)$, and the stochastic process $(w_t, R_t, z_t^i)_{t=1}^{\infty}$ has distribution $P^i(\cdot)$ $(= P^i(\cdot; w_0, R_0, z_0^i, \theta^M))$. We follow Epstein and Zin (1989) in defining utility directly on these random consumption streams, so given the savings decision $a^{i,\prime}$ utility of the date 0 self equals $U^{\epsilon_0^i,i}(c_0, c_1, c_2, \ldots)$ where $U^{\epsilon_0^i,i}$ is the household's utility function and c_0, c_1, c_2, $\ldots$ the random consumption stream just defined. As in Epstein and Zin (1989), $U^{\epsilon_0^i,i}$ need not have an expected utility representation; but if it does $U^{\epsilon_0^i,i}(c_0, c_1, c_2, \ldots) = \int u^{\epsilon_0^i,i}(\hat{c}_0, \hat{c}_1, \hat{c}_2, \ldots)\ P^{\theta,i}_{w_0,R_0,z_0^i}(d(w_t, R_t, z_t^i)_{t=1}^{\infty})$, where $\hat{c}_t$ equals c_t conditioned on the realization at date t, and $u^{\epsilon_0^i,i}(\hat{c}_0, \hat{c}_1, \hat{c}_2, \ldots) = U^{\epsilon_0^i,i}(\delta_{\hat{c}_0}, \delta_{\hat{c}_1}, \delta_{\hat{c}_2}, \ldots)$ with δ_x denoting a degenerate distribution with unit mass as x. Note that through the random utility parameter ϵ_t^i, a "current self" may be uncertain about the behavior of "future selves" (see Example 4 below). Note finally that, unlike for example Epstein and Zin (1989), we do not assume that U^{ϵ_0} is time stationary/dynamically consistent (recursive), and for example discounting could be quasi-geometric as in Laibson (1997) (Example 1).

The previous description is quite general and the key behavioral restrictions our formulation imposes is the time-stationarity (also called time invariance; see Halevy (2015)) of both the belief

[33]For example, the household may expect the (future) interest rate to be equal to $\sigma_t^i R_t$, where R_t is the (perfect foresight after-tax) interest rate and σ_t^i is *i.i.d.* random noise with mean $\gamma \in \mathbb{R}_+$ and captures the extent of uncertainty faced by the agent. If $\gamma = 1$, the household is on average correct about the interest rate but is not able to predict it perfectly; if $\gamma \neq 1$, the household would be making systematic mistakes in predicting the interest rate.

formation process and utility. The justification would be that our description begins at some date where actual beliefs and utilities have converged to time-independent limits (see Section 3.3 and Section 4 for detailed discussion of this approach's limitations, in particular, which types of beliefs formations are ruled out). It then follows that, conditioned on initial conditions (a_t^i, w_t, R_t, and z_t^i), the decision problem of the "current self" at date t is the same as the decision problem any "future self" would face if it were to face the (exact) same initial conditions.[34] We impose in the next definition such *time-stationarity* formally by defining the savings function of a household as a (symmetric) Nash equilibrium of the Bayesian game between future selves and the current self (time-consistency is then a special case). Note that under time-stationarity, there is no loss of generality in focusing on the self at date $t = 0$ as in the description above.

Let $\underline{a}^i \in \mathbb{R}$ denote household i's *borrowing constraint* which may be explicit and occasionally binding as in Aiyagari (1994), or may be derived from a more fundamental transversality condition (see Aiyagari (1994), pp.665-666). The upper bound $\overline{a}^i \in \mathbb{R}$ comes with no loss of generality within the general setting of this paper since it may be chosen so that it never binds in equilibrium (P^i-almost surely and for almost every agent) under effective compactness in production (see Section 3.2).

Definition 6 ***(Time-Stationary Savings Functions and the Savings Correspondence)*** *$s^i_{w,R,z^i} : \mathbb{R} \to \mathbb{R}$ is a time-stationary savings function (TSSF) if for all a^i, w, R, and a.e. $z^i = (l^i, \epsilon^i, \sigma^i)$, and conditioned the random sequence $(w_t, R_t, z_t^i)_{t=1}^{\infty}$ with distribution $P^i(\cdot) = P^i(\cdot; w, R, z^i, \theta^M)$:*

$$s^i_{w,R,z^i}(a) \in \arg\max_{a' \in [\underline{a}^i, \bar{a}^i] : a' \leq (1+R)a^i + wl^i} U^{i,\epsilon^i}((1+R)a^i + wl^i - a', (1+R_1)a' + w_1 l_1^i - s^i_{w_1,R_1,z_1^i}(a'), (1+R_2)s^i_{w_1,R_1,z_1^i}(a') + w_2 l_2^i - s^i_{w_2,R_2,z_2^i}(s^i_{w_1,R_1,z_1^i}(a')), \ldots)] \, . \tag{18}$$

The savings correspondence $S^{\theta,i}_{w,R,z^i} : \mathbb{R} \to 2^{\mathbb{R}}$ is defined as the union of all time-stationary savings functions.

As a final remark, we have tried to strike a balance between generality and expositional simplicity here; in particular, we have assumed that the households can be ascribed specific beliefs (possibly non-additive as in the case of ambiguity).[35]

[34]Note that without time-stationarity, savings functions would not be time invariant and steady states would not exist.

[35]With set-valued beliefs that cannot be discriminated amongst, *i.e.*, when P^{θ,i^M}_{w,R,z^i} is a set of measures (see also footnote 18), there will be "multiple versions" of an individual associated with each of the possible beliefs. One then determines the set of TSSFs for each belief in the set of beliefs (each "version" of the household) and takes the union as above to obtain the savings correspondence. Since we already allow savings to be set-valued, this poses no additional problems for any of our results as long as the set of beliefs as a correspondence varies upper hemi-continuously with w, R, and $z^i$and has a compact range.

As in Epstein and Zin (1989), all sets are equipped with the Borel σ-algebra and the topology on probability measures and on random sequences is the weak convergence topology (Epstein and Zin (1989), p.940). Throughout the following we impose very weak continuity conditions on utility and beliefs, as well as compactness on the set of random utility parameters.

Condition C *$U^{i,\epsilon^i}(c_0, c_1, c_2, \ldots)$ is continuous in $(\epsilon^i, c_0, c_1, \ldots)$, $P^i(B; w, R, z^i, \theta^M)$ is continuous in (w, R, z) for any measurable set B, and $\epsilon^i \in E^i$ where $E^i \subseteq \mathbb{R}$ is compact.*

6.2 Models With Dynamic Inconsistency

From now on we omit the household index i to simplify notation. We begin with the case where the objective in (6) is not required to be dynamically consistent. As explained previously, we limit attention to weakly separable utilities (*i.e.*, utility functions satisfying (19) below). Recall that for a measure P on a set B and a measurable function $U : B \to \mathbb{R}$, a certainty equivalent (also known as a generalized mean) is a function of the type $\mu_P[U] = g^{-1}\left(\int g(U(b))P(db)\right)$, where g is strictly increasing. The integral is here the Lebesgue integral if the measure P is additive, and the Choquet integral if P is non-additive. We then have:

Proposition 4 *Assume that*

$$U^\epsilon(c_0, c_1, c_2, \ldots) = H\left(u_0^\epsilon(c_0) + \beta h\left(\mu_{P(\cdot;w,R,z,\theta^M)}[\tilde{U}^\epsilon(c_1, c_2, \ldots)]\right)\right), \tag{19}$$

where H and h are strictly increasing functions, u_0^ϵ is concave for all ϵ, $\mu_{P(\cdot;w,R,z,\theta^M)}$ is a certainty equivalent based on the household's beliefs conditioned on w, R, z and θ^M, and $\beta > 0$ is a positive constant ("patience"). Then if Condition C holds, the savings correspondence $S^\theta_{w,R,z}(a)$ satisfies Assumption 1.

As with all other results in this section, the proof is relegated to Appendix B. We next show that several leading cases of behavioral preferences are covered by this proposition.

Example 1 ***(Hyperbolic, Quasi-Hyperbolic, and General Delay Discounting)*** *Deterministic models with non-geometric discounting were introduced in Section 3.1 (see 3.1.1). Under perfect foresight, we can suppress beliefs entirely (by setting all variables on the right-hand side of (18) equal to their actual values); and it is clear that any such model of delay discounting is then a special case of (4) by taking H, h, and the certainty equivalent equal to the identity function. These models are therefore covered by Proposition 4, and Assumption 1 is satisfied if u_0 is continuous and concave. Under hyperbolic discounting $\tilde{U}^\epsilon(c_1, c_2, \ldots) = \sum_{t=1}^\infty (1 + \alpha t)^{-\frac{\gamma}{\alpha}} u_0(c_t)$ (Loewenstein and Prelec (1992)), and under quasi-hyperbolic discounting, $\tilde{U}^\epsilon(c_1, c_2, \ldots) = \beta \sum_{t=1}^\infty \delta^t u_0(c_t)$ (Laibson (1997)). More generally,*

$\tilde{U}^{\epsilon}(c_1, c_2, \ldots) = \sum_{t=1}^{\infty} f(t)u_0(c_t)$ where $f(t)$ is date t discounting. These are all admissible within our general framework because we do not insist that the agent's decision problem can be cast as a dynamic programming problem. Note also that the extension to random environments and/or general beliefs is straightforward in light of the general specification in (4).

Example 2 ***(Finite Planning Horizons, Sparse Maximization)*** *A deterministic, perfect foresight model with a finite planning horizon, $U^{\epsilon}(c_1, c_2, \ldots) = \sum_{t=1}^{T} \beta^t u(c_t)$, immediately fits into (19) and so Assumption 1 is satisfied if u is continuous and concave. Extensions to random environments (possibly with ambiguity) are straightforward and can also be shown to satisfy Assumption 1 under Condition C. As mentioned in Section 3.1.3, the finite planning horizon model may be viewed as a particularly simple (reduced-form) expression of sparsity in the sense of Gabaix (2014, 2017). Richer forms of sparsity constraints, for example, in the form of additional restrictions on the set of choice variables, can also be incorporated into our setup. A particularly fruitful approach is to replace the* $\max$ *operator in (18), with the "sparse* $\max$*" operator of Gabaix. As explained in Gabaix (2017), the "sparse* $\max$*" formulation is quite tractable and also implies a "sparse" Bellman operator which is a monotone contraction (see Gabaix (2017), Lemma 3.6). General savings correspondences for sparse maximization can then be derived from this formulation and naturally satisfy Assumption 1 (which can be proved formally from slight modifications of Proposition 4).*

Example 3 ***(Systematically Wrong Beliefs)*** *Imagine that a household systematically (i.e., period after period) forms incorrect beliefs, or misperceives a key economic variable such as the relevant interest rate (the behavioral agent considered in Section 5.3 falls into this category and can therefore be formally microfounded along the lines described next). As a simple illustration that also fits into (19), imagine that (objectively) the world is deterministic, and that the household has perfect foresight with respect to all variables except for the interest rate which it systematically overestimates. Specifically, at any date t, the beliefs P are such that the household expects the interest rate at any future period t' to equal $R_{t'} + \hat{\sigma} > R_{t'}$ (note that in this case $z = (\epsilon, l, \sigma)$, where $\sigma = 0$, is fixed/deterministic in the "true model", and the agents expectations of z, $(\epsilon, l, \hat{\sigma})$ are deterministic too). To simplify, assume also that the utility objective is additive with geometric discounting. By Definition 6, the agent's TSSF is then determined by the requirement that for all a and for $\sigma \in \{0, \hat{\sigma}\}$:*

$$\begin{aligned} s_{w,R,\epsilon,l,\sigma}(a) \in \quad & \arg\max_{a' : (1+R+\sigma)a + wl \geq a'} u^{\epsilon}\left((1+R+\sigma)a + wl - a'\right) + \beta u^{\epsilon}\left((1+R+\hat{\sigma})a' + \right. \\ & \left. wl - s_{w,R,\epsilon,l,\hat{\sigma}}(a')\right) + \beta^2 u^{\epsilon}\left((1+R+\hat{\sigma}) s_{w,R,\epsilon,l,\hat{\sigma}}(a') \right. \\ & \left. + wl - s_{w,R,\epsilon,l,\hat{\sigma}}(s_{w,R,\epsilon,l,\hat{\sigma}}(a'))\right) + \ldots . \end{aligned} \tag{20}$$

To find the TSSF, first solve (20) with $\sigma = \hat{\sigma}$ in order to determine what may be called "the misperceived savings function" of the future selves, $s_{w,R,\epsilon,l,\hat{\sigma}}$ (note that this problem is recursive and it can therefore *be solved by standard dynamic programming techniques). Then solve (20) for $\sigma = 0$ given $s_{w,R,\epsilon,l,\hat{\sigma}}$ to obtain the actual savings function, $s_{w,R,\epsilon,l,\sigma}(a)$. This same two-step procedure generalizes to arbitrary misperceptions/systematically wrong beliefs: First, solve a standard (possibly stochastic) dynamic programming problem given the household's (wrong) beliefs to find the misperceived savings function of the future selves; then solve the current self's optimization problem given this misperceived savings function.*

Note that in this example, a perfect commitment solution would not satisfy the budget identities at $t = 1, 2, \ldots$, and it is therefore inadmissible. Thus dynamic inconsistency is embedded in the beliefs/misperceptions.

While it seems implausible that an individual should period after period suffer from the same misconception as in the example just given, more realistic cases of misperceptions similarly fit into this framework. We already presented a more realistic instance of savings behavior resulting from this type of microfoundation in Section 5.3. For a second example, imagine that the agent's beliefs about σ above is not deterministic but stochastic and has the correct mean 0 and non-zero variance. This would be a type of "anxiety" (worrying about the future when in reality, i.e., in the true model, there is actually nothing to worry about), and if the TSSF is increasing and convex in savings, it would make the household save more than under correct beliefs. Economically, this behavioral model thus predicts savings above what the pure precautionary motive predicts (see Jensen (2018) on convexity of savings functions and the relationship with precautionary savings).

In the next subsection we consider a number of additional examples (such as ambiguity) that fit into the dynamically consistent case. But each of these might in addition also feature dynamic inconsistency. If so, the examples would still be covered by Proposition 4 if the objective is weakly separable.

6.3 Dynamically Consistent Recursive Models

We next turn to the second part of Assumption 1 (the savings correspondence increasing in a) in the case where preferences are dynamically consistent. Here we provide results for the most general recursive preferences in Epstein and Zin (1989), which as a special case include Kreps and Porteus (1978) and in the deterministic case reduces to the class of recursive utilities (Koopmans (1960)).[36] The certainty equivalent mentioned in the proposition was defined in the previous subsection. Supermodularity of W has the usual meaning (*e.g.*, Topkis (1978)) and is equivalent to assuming that consumption at different dates are Edgeworth-Pareto complements (Chipman (1977)). Again the proof has been placed in Appendix B.

[36]For a survey covering both cases see Backus, Routledge and Zin (2004).

Proposition 5 *Suppose that*

$$U^{\epsilon}(c_0, c_1, c_2, \ldots) = W(u^{\epsilon}(c_0), \beta\mu_{P(\cdot;w,R,z,\theta^M)}[U^{\epsilon'}(c_1, c_2, \ldots)]) \,, \tag{21}$$

where $\mu_{P(\cdot;w,R,z,\theta^M)}$ is a certainty equivalent based on the household's beliefs conditioned on w, R, z and θ^M, u^{ϵ} is concave and increasing, the time-aggregator $W(u, U)$ is concave in u, and increasing and supermodular in (u, U), and $\beta > 0$ is a positive constant ("patience"). Then if Condition C holds, the savings correspondence $S^{\theta}_{w,R,z}(a)$ satisfies Assumption 1.

We next discuss several applications of this proposition.

Example 4 ***(Random Utility, Mistakes, Approximate Rational, Satisficing Behavior, Over-optimism, and Quantal Response Equilibrium)*** *The random utility models introduced in Section 3.1.2 fit straightforwardly into (21) if the objective is dynamically consistent, and into (19) if it is not. This is the case, for example, when beliefs $P(\cdot) = P(\cdot; w, R, z, \theta^M)$ are probabilistically correct given the model θ^M (i.e., under rational expectations). But Proposition 5 continues to apply beyond the simple case and ensures that Assumption 1 holds under natural continuity, concavity, and in the case of (21), complementarity conditions (these ensure that goods are normal as explained in Chipman (1977)). To expand on Section 3.1.2, imagine that ϵ_t (objectively, i.e., as expressed through μ_z) is i.i.d. with mean 0 but that the agent through $P(\cdot; w, R, z, \theta^M)$ (incorrectly) believes that ϵ_t first-order stochastically dominates the objective probability (and in particular has mean greater than 0). If savings increases in ϵ, the agent is then "overly optimistic" about his future frugality which causes him too save less today than he would if his beliefs were correct. At the following date, the agent will of course be "disappointed" for not living up to his own expectations — but with a TSSF he goes on to assume that next year he will start saving more.*[37] *If instead ϵ parametrizes selves' subjective beliefs, a TSSF is the quantal response equilibrium (McKelvey and Palfrey (1995)) of the game the current self plays with future selves.*

Example 5 ***(Ambiguity)*** *If a household has incomplete information about the objective probabilities governing the random disturbances in z, then even under rational expectations ($P(\cdot) = P(\cdot; w, R, z, \theta^M)$ probabilistically correct given the model θ^M and current observations) it cannot be attributed unique subjective beliefs unless it satisfies the axioms of Savage (1954) (note that this has nothing to do with whether the subjective beliefs are right or wrong). Since we have allowed for beliefs $P(\cdot)$ to be non-additive, in which case the objective in (18) is the agent's Choquet Expected Utility (CEU), most models of ambiguity are immediately covered by the previous results. In particular, non-additive $P(\cdot)$ may reflect "multiple priors"since CEU with convex capacities equals the minimum expected utility over the probabilities in the*

[37] This situation also leads to dynamic inconsistency, and in that case, we have to assume weak separability in order to apply Proposition 4.

capacity's core (Schmeidler (1989)). Note that ambiguity is also covered in the dynamically inconsistent case of the previous subsection provided that the underlying utility function is weakly separable.

Example 6 ***(Epstein-Zin and Kreps-Porteus Preferences)*** *Equation (21) corresponds to the general recursive specification of Epstein and Zin (1989). Epstein and Zin's focus is on the case where W is a CES function (Epstein and Zin (1989), p. 946). If, in addition, the certainty equivalent is just the mean of $(U(\cdot))^{\alpha}$ raised to the power $\frac{1}{\alpha}$, this yields Kreps-Porteus preferences (Kreps and Porteus (1978), see also Epstein and Zin (1989), p. 947-948):*

$$U(c_0, m) = \left[c_0^{\rho} + \beta\left(\mathbb{E}_m[(U(\cdot))^{\alpha}]\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} .$$

The conditions in Proposition 5 hold under the assumptions imposed by Epstein and Zin (in particular, the CES aggregator satisfies the proposition's conditions). Hence Assumption 1 will be satisfied by the implied savings correspondence.

Example 7 ***(Self-Control and Temptation)*** *We have so far focused on the Epstein-Zin type formulation of dynamic choices where preferences are defined over random consumption streams. Gul and Pesendorfer (2001, 2004), instead, model temptation and self-control by defining preferences directly on choice problems which are combinations of decisions today and the resulting continuation problem. Specifically, consider as above an agent with assets a at date 0 and therefore income $(1 + R_0)a + w_0 l_0^i$. If the agent saves a_1 (and so consumes $(1+R_0)a+w_0 l_0^i - a_1$), this implies (random) income $(1+R_1)a_1+w_1 l_1^i$ at the following date, and so recursively a new choice problem (a continuation problem in the language of Gul and Pesendorfer). As discussed in Gul and Pesendorfer (2004), Section 6, this formulation is mathematically simpler than hyperbolic models of preference reversal because it implies unique optimal payoffs (so, as in our general dynamically consistent case, there is no need to resolve multiplicity by considering game-theoretic interaction between multiple selves). In particular, this approach leads to well-defined recursive programs and therefore the resulting savings correspondences satisfy Assumption 1 under continuity and compactness conditions on primitives (exactly as in our general formulation above) and are increasing in assets provided that the fundamental utilities are concave.*

6.4 The Effects of Changes in Patience

In this subsection, we show that when a household becomes more patient, *i.e.*, when β increases in either of the specifications covered in the previous two subsections, then savings increase. It then follows from Proposition 2 that a change to a more patient environment — meaning that a subset of households become more patient while the rest do not change their preferences —

is a local positive shock.[38] The cases with dynamic consistency and dynamic inconsistency are covered separately.

Proposition 6 *If a household satisfies the conditions of Proposition 4 with U concave and continuously differentiable, and the optimal strategies are linear in assets, then a change to a more patient environment is a local positive shock at any $k' \geq 0$.*

Since savings will equal zero if assets $a = 0$, the only smooth class of additive utility functions that lead to linear strategies is the isoelastic one (*i.e.*, the one where period utility function equal to either $\alpha \log x$ or $\frac{1}{1-\rho} x^{1-\rho}$, $\rho > 0$, see Pollak (1971), p. 402). If U is not additive, a sufficient condition is that u_0 and $\tilde{U}$ are homogenous of the same degree.

Proposition 7 *If a household satisfies the conditions of Proposition 5 and has additive subjective beliefs, then a change to a more patient environment is a local positive shock at any $k' \geq 0$.*

7 Concluding Remarks and Future Directions

A common conjecture is that equilibrium analysis becomes excessively challenging in the presence of behavioral preferences and biases, thus implicitly justifying a focus on models with time-additive, dynamically consistent preferences and rational expectations. In this paper, we demonstrated that, in the context of the behavioral neoclassical growth model — the one-sector neoclassical growth model enriched with the large class of behavioral preferences — this conjecture is not necessarily correct. Results concerning the direction of change in the long run (or "robust comparative statics" for the steady-state equilibrium) can be obtained for a wide range of behavioral preferences and rich heterogeneity. Put simply, our main results state the following: for any change in policy or underlying production or preference parameters of the model, we first determine whether this is a *local positive shock*; this step involves no equilibrium analysis, but only the determination of whether at *given* prices (and thus given the initial capital-labor ratio), there will be greater savings. The emphasis on "local" in local positive shock is precisely to underscore that all of this is at given prices, and there is no presumption or necessity that such a shock will increase savings at other prices or capital-labor ratios. Then under mild regularity conditions (satisfied for all behavioral preferences we have discussed in this paper), no matter how complex the equilibrium responses are, they will not overturn the direction of the initial change and thus the steady-state equilibrium will involve a greater capital-labor ratio (and the

[38] As usual, some households might lower their savings as long as there are enough households to ensure that on average, the direct effect is positive (Section 5.1).

changes in prices that this brings). Conversely, if the initial change is a *local negative shock*, then the long-run capital-labor ratio will decline. No further information than whether a change in policy or parameters is a local positive or negative shock (and the verification of the mild regularity conditions, which we have already established for a range of behavioral models) is necessary for these conclusions.

At the root of this result is a simple and intuitive observation: in the one-sector model, the only way the direction of the impact of the initial impetus (say a local positive shock) can be reversed is by having the equilibrium response to this initial shock to go strongly in the opposite direction. For example, savings could decline strongly in response to a higher capital-labor ratio. But either such an equilibrium response would still not overturn the initial local positive shock, in which case the conclusion about the steady-state equilibrium applies. Or it would overturn it and reduce the long-run capital-labor ratio, but in this case the perverse effect would go in the direction of strengthening, not reversing, the initial local positive shock.

This intuition also clarifies the limitations of our results. A similar logic would not apply if the economy had multiple state variables rather than the single state variable as in our (one-sector) behavioral neoclassical growth model. In such richer circumstances, similar results would necessitate supermodularity conditions for the set of state variables or a result that in the relevant problems the vector of state variables could be reduced to be functions of a single overall state variable. One example in which this latter approach can be used straightforwardly is an extension of our setup to a multi-sector neoclassical growth model. For brevity, we did not develop the details of this model, but the main idea is simple. Suppose that we have a n-sector growth model with no irreversibilities, neoclassical production functions in each sector and competitive capital markets (though distortions that differ across sectors can be introduced for additional generality). Then the marginal return to capital has to be equalized across different sectors, which determines an allocation of the overall capital stock across sectors and enables us to have a reduced-form problem just as a function of the overall capital stock. Then similar comparative static results can be developed for this overall capital stock in this type of multi-sector environment. Beyond this case, extending our results to other settings with multiple state variables is far from trivial, and would typically necessitate strong supermodularity/monotonicity conditions (in contrast, our current results require no such monotonicity assumptions).

One obvious limitation of our approach bears repeating at this point: our focus has been on comparative statics, and thus on qualitative rather than quantitative results. Many questions in modern macroeconomics necessitate quantitative analysis, and the quantitative impact of a policy change may critically depend on behavioral biases and the exact structure of preferences even if the direction of long-run change does not. An obvious but challenging area for future research is to investigate when certain quantitative conclusions may not depend on appropriately introduced behavioral biases or heterogeneity (for example, in the sense that as behavioral assumptions are changed, quantitative change in some key variables remains near changes implied by a benchmark model).

We should also again emphasize that our results should not be read as implying that behavioral biases and deviations from the benchmark model of time-additive, dynamically consistent preferences and rational expectations are unimportant. What we have established is that they do not change the direction of long-run responses in the one-sector neoclassical growth model provided that they do not alter the direction of the initial impulse. But we have also demonstrated via examples how behavioral considerations can easily turn a change in policy or parameters that would have otherwise been a local positive shock into a local negative shock. Then our result works in reverse: no matter what the equilibrium responses are, this impact of behavioral considerations cannot be reversed and the long-run response of the economy will be the opposite of the response of an economy inhabited by households with standard preferences and rational expectations. In this instance, therefore, the power of behavioral biases and richer preferences to impact macroeconomic equilibrium outcomes is amplified. In this light, another important and challenging area is to characterize in greater detail what types of realistic behavioral considerations, and under what circumstances, will change the direction of initial changes in policy or parameters from a local positive to a local negative shock.

Appendix A. Changes in the Environment: A Topological Approach, Discussion of Related Literature

Since this section's observations may be of independent interest and apply not only to market correspondences, we are going to view the market correspondence $\mathcal{M} : K \times \Theta \to 2^{\mathbb{R}}$, $K \subseteq \mathbb{R}$, more abstractly and impose any necessary assumptions directly. Denote by $m_S^\theta(k) = \inf \mathcal{M}^\theta(k)$ and $m_L^\theta(k) = \sup \mathcal{M}^\theta(k)$ the least and greatest selections, and by $k_S^\theta = \inf\{k \in K : k \in \mathcal{M}^\theta(k)\}$ and $k_L^\theta = \sup\{k \in K : k \in \mathcal{M}^\theta(k)\}$ the least and greatest fixed points (when they exist, which of course they do if $\mathcal{M}$ is a market correspondence). Now equip Θ with an order as well as a topology (in the simplest situation where we consider a change in just a single parameter, Θ may be taken to be a subset of $\mathbb{R}$, and these would therefore be the usual/Euclidean order and topology, respectively). A function $m : \Theta \to \mathbb{R}$ is said to be (i) *increasing* if $\theta \leq \hat{\theta} \Rightarrow m(\theta) \leq m(\hat{\theta})$ for all $\theta, \hat{\theta} \in \Theta$, and (ii) *locally increasing at* $\theta^* \in \Theta$ if $\theta \leq \hat{\theta} \Rightarrow m(\theta) \leq m(\hat{\theta})$ for all $\theta, \hat{\theta}$ in an open neighborhood of θ^*. Finally, say that $\mathcal{M}$ *begins above and ends below the 45° line* if $m_*(\inf K, \theta) \geq \inf K$ and $m^*(\sup K, \theta) \leq \sup K$.

Theorem 5 ***(Abstract Shifts in Fixed Point Correspondences)*** *Consider an upper hemi-continuous and convex valued correspondence $\mathcal{M} : K \times \Theta \to 2^{\mathbb{R}}$ where K is a compact subset of $\mathbb{R}$ and Θ is a compact subset of an ordered topological space. Assume also that the graph begins above and ends below the 45° line for all $\theta \in \Theta$. Then the least and greatest fixed points k_S^θ and k_L^θ are increasing in θ if for all $\theta^* \in \Theta$, $m_L^\theta(k_L^{\theta^*})$ and $m_S^\theta(k_S^{\theta^*})$ are locally increasing in θ at θ^*.*

Proof. Consider the greatest fixed point $k_L^{\theta^*}$ given some $\theta^* \in \Theta$. To simplify notation, we take $\Theta \subseteq \mathbb{R}$ (but the argument is true in general). Since $m_L^\theta(k_L^{\theta^*}) \geq m_L^{\theta^*}(k_L^{\theta^*}) = k_L^{\theta^*}$ for $\theta^* + \epsilon > \theta > \theta^*$, $m_L^\theta(\cdot)$ begins above the 45° line and ends below it on the interval $[k_L^{\theta^*}, \sup K]$. Since $\mathcal{M}$ has convex values, $\mathcal{M}^\theta(\cdot)$ therefore has a fixed point on this interval, and so $k_L^\theta \geq k_L^{\theta^*}$. This argument clearly extends to any $\theta > \theta^*$ (not necessarily in a neighborhood) since we may reach any such θ in a finite number of steps (Θ is compact so any open cover contains a finite subcover). The more difficult case is when $\theta^* - \epsilon < \theta < \theta^*$. Assume for a contradiction that $k_L^\theta > k_L^{\theta^*}$. Consider θ_n, where $\theta < \theta_n < \theta^*$. Since $\theta_n > \theta$, it follows from the first part of the proof that $k_L^{\theta_n} \geq k_L^\theta > k_L^{\theta^*}$. Note that these inequalities hold for any $\theta_n \in (\theta, \theta^*)$. Since K is compact, we may consider a sequence $n = 0, 1, 2, \ldots$ with $\theta_n \uparrow \theta^*$ and such that $\lim_{n\to\infty} k^*(\theta_n)$ exists. $k_L^{\theta_n} \in \mathcal{M}^{\theta_n}(k_L^{\theta_n})$ for all n and $\mathcal{M}$ has a closed graph, hence $\lim_{n\to\infty} k_L^{\theta_n} \in \mathcal{M}^{\theta^*}(\lim_{n\to\infty} k_L^{\theta_n})$. But since $\lim_{n\to\infty} \lim_{n\to\infty} k_L^{\theta_n} \geq k_L^\theta > k_L^{\theta^*}$, this contradicts that $k_L^{\theta^*}$ is the greatest fixed point. The parallel statement for the least fixed point $k_S^{\theta^*}$ is shown by a dual argument (in this case the situation where $\theta^* - \epsilon < \theta < \theta^*$ is simple while the limit sequence argument must be used for the case where $\theta^* + \epsilon > \theta > \theta^*$). ■

Corollary 1 ***(Local Positive Shocks, Topological Case)*** *Let the assumptions of Theorem 1 hold and assume in addition that Θ is a compact subset of an ordered topological space and that the market correspondence $\mathcal{M}^\theta(k)$ is upper hemi-continuous in (θ, k). Then the greatest and least steady states k_S^θ and k_L^θ are increasing in θ if for all $\theta^* \in \Theta$ and all $\theta_a < \theta_b$ in a neighborhood of θ^*, the change in the environment from θ_a to θ_b is a local positive shock at $k_L^{\theta^*}$ as well as a local positive shock at $k_S^{\theta^*}$.*

It is useful at this point to briefly contrast the main results of our paper, including those presented in this Appendix, to other equilibrium comparative static results in the literature. Most of the results in the literature are similar to those of Milgrom and Roberts (1994) who show that when the equivalent of our market correspondence $\mathcal{M}$ is "continuous but for jumps up" and its graph shifts up (meaning that $m_L^\theta(k)$ and $m_S^\theta(k)$ are increasing in θ *for all* k), then the least and the greatest fixed points increase (see, for example, their Corollary 2).[39] Let us refer to this well-known result as the "for all k curve shifting theorem". Theorem 5 is very different from this result. It shows instead that if $\mathcal{M}$ is upper hemi-continuous in (k, θ) (rather than just in k, *cfr.* footnote 39), the same conclusion requires only that the correspondence shifts up at the least and the greatest fixed points, k_S^θ and k_L^θ. The results presented in Section 4 similarly require only local shifts in steady states. That we only need to verify that $\mathcal{M}$ shifts up *locally*, in particular, *at* the steady states, is the key technical contribution of the paper and plays a critical role for all of our our results.[40]

To explain a little further, let us consider a particularly simple case where a dynamic economy can be reduced to a fundamental equation of the form

$$G(k_t, k_{t-1}, \theta) = 0 \, , \tag{22}$$

where $\theta \in \mathbb{R}$ is an exogenous parameter, $k_t \in \mathbb{R}$ is capital, or the capital-labor ratio, at date t and $G : \mathbb{R}^3 \to \mathbb{R}$ a smooth function. In this case, the market correspondence can be defined as

$$\mathcal{M}^\theta(k) = \{\hat{k} : G(\hat{k}, k, \theta) = 0\} \, . \tag{23}$$

In the Ramsey-Cass-Koopmans model, for example, $G(k_t, k_{t-1}, \theta) = 0 \Leftrightarrow k_t = g(k_{t-1}, \theta)$, and then $\mathcal{M}^\theta(k) = g(k, \theta)$. Clearly, k^* is a steady state given θ if and only if $k^* \in \mathcal{M}^\theta(k^*)$. Note,

[39] $\mathcal{M}$ is *continuous but for jumps up* if it has convex values, $\limsup_{x^n \uparrow x} m^*(x^n, t) \leq m_L^\theta(k)$, and $\liminf_{x^n \downarrow x} m_*(x^n, t) \geq m_S^\theta(k)$.

Acemoglu and Jensen (2013) shows that if $\mathcal{M}$ is upper hemi-continuous in k and has convex values, then it is continuous but for jumps up.

[40] The other important technical ingredient is the definition of the market correspondence $\mathcal{M}$ and our Lemma 1, which enables us to work with a simple, albeit abstract, mapping.

however, that (22) — even in the more general form $0 \in G(k_t, k_{t-1}, \theta)$ where G is a correspondence — is not general enough to nest our behavioral neoclassical growth model (because we also need to condition on the distribution of assets). Nevertheless, (22) is useful to provide the technical intuition for our main results since both in the case of (23) and our Definition 2, the market correspondence is constructed by conditioning on the information that the capital-labor ratio in question, k, has to be consistent with a steady-state equilibrium. In particular, the fact that, with the conditioning on the steady state k^*, (23) a one-dimensional fixed point problem allows us to use "curve shifting" arguments without imposing any type of monotonicity on the dynamical system defined by (22) (see also Acemoglu and Jensen (2015) for a related discussion of non-monotone methods). Given $\mathcal{M}^\theta(k)$ and this construction, Theorem 5 and the results presented in Section 4 enable us to predict how the greatest and the least steady states vary with θ when $\mathcal{M}^\theta(k)$ shifts up locally starting at these steady states (and provided that $\mathcal{M}$ satisfies the relevant theorem's regularity conditions).

The added generality and flexibility is considerable. In many applications, including the problem of equilibrium analysis in the behavioral neoclassical growth model we focus on in this paper, the conditions for the "for all k curve shifting theorem" will not hold even if (22) applies. This is for both substantive and technical reasons. Substantively, in economies such as the behavioral neoclassical growth model the possible heterogeneity in the responses of agents to changes in the environment would often preclude such uniform shifts. To see the technical problem, suppose that we were checking these conditions using the implicit function theorem. That would amount to verifying that $\frac{dk}{d\theta} > 0$ for all $\tilde{k}$ while $G(k, \tilde{k}, \theta) = 0$ holds. But since the implicit function theorem requires as a minimum that $D_k G(k, \tilde{k}, \theta) \neq 0$, and "running through all $\tilde{k}$'s" will almost invariably violate this condition for some $\tilde{k}$, this method will generally fail (order theoretic methods are of no help here either; and of course, it is *not* enough to show that $\frac{dk}{d\theta} > 0$ for almost every $\tilde{k}$ because any point we fail to check may precisely be a point where the market correspondence "jumps"). When we only need to check local conditions, these difficulties are bypassed.

References

Acemoglu, D. (2009), *Introduction to Modern Economic Growth*, Princeton University Press, 2009.

Acemoglu, D. and M.K. Jensen (2010): "Robust Comparative Statics in Large Static Games", *IEEE Proceedings on Decision and Control* 49.

Acemoglu, D. and M.K. Jensen (2013): "Aggregate Comparative Statics", *Games and Economic Behavior* 81, 27-49.

Acemoglu, D. and M.K. Jensen (2015): "Robust Comparative Statics in Large Dynamic Economies", *Journal of Political Economy* 123, 587-640.

Ahn, H., G. Kaplan, B. Moll, T. Winberry and C. Wolf (2018): "When Inequality Matters for Macro and Macro Matters for Inequality", Working Paper, Princeton University.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2018): "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach", *NBER Macroeconomics Annual*.

Aiyagari, S.R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving", *Quarterly Journal of Economics* 109, 659-684.

Al-Najjar, N.I. (2004): "Aggregation and the Law of Large Numbers in Large Economies", *Games and Economic Behavior* 47, 1-35.

Aubin, J.-P. and H. Frankowska, *Set-Valued Analysis*, Birkhauser, Boston, 1990.

Aumann, R.J. (1965): "Integrals of Set-Valued Functions", *Journal of Mathematical Analysis and Applications* 12, 1-12.

Backus, D.K., B.R. Routledge, and S.E. Zin (2004): "Exotic Preferences for Macroeconomists", *NBER Macroeconomics Annual* 19, 319-390.

Balbus, L., K. Reffett, and L. Wozny (2015): "Time Consistent Markov Policies in Dynamic Economies with Quasi-hyperbolic Consumers", *International Journal of Game Theory* 44, 83-112.

Barro, R.J. (1999): "Ramsey Meets Laibson in the Neoclassical Growth Model", *Quarterly Journal of Economics* 114, 1125-1152.

Barro, R.J. and X.I. Sala-i-Martin, *Economic Growth*, MIT Press, 2004. 2

Becker, G.S. (1962): "Irrational Behavior and Economic Theory", *Journal of Political Economy* 70, 1-13.

Becker, R.A. and J.H. Boyd, *Capital Theory, Equilibrium Analysis, and Recursive Utility*, Blackwell, 1997.

Beshears J., J.J. Choi, A. Fuster, D. Laibson and B.C. Madrian (2013): "What Goes Up Must Come Down? Experimental Evidence on Intuitive Forecasting", *American Economic Review* 103, 570-574.

Cagan, P. (1956): "The Monetary Dynamics of Hyperinflation", In Friedman, M. (ed.). Studies in the Quantity Theory of Money. Chicago: University of Chicago Press.

Cao, D. and I. Werning (2017): "Saving and Dissaving with Hyperbolic Discounting", *Econometrica*, forthcoming.

Chipman, J.S. (1977): "An Empirical Implication of Auspitz-Lieben-Edgeworth-Pareto Complementarity", *Journal of Economic Theory* 14, 228-231.

Epstein, L.G. and J.A. Hynes (1983): "The Rate of Time Preference and Dynamic Economic Analysis", *Journal of Political Economy* 91, 611-635.

Epstein, L.G. and S.E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica* 57, 937-969.

Epstein, L.G. and S.E. Zin (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis", *Journal of Political Economy* 99, 263-286.

Fudenberg, D. and D.K. Levine (2006): "A Duel Self Model of Impulse Control", *American Economic Review* 96, 1449-1476.

Fudenberg, D. and D.K. Levine (2012): "Timing and Self-Control", *Econometrica* 80, 1-42.

Fuster A., B. Hebert and D. Laibson (2012): "Investment Dynamics with Natural Expectations", *International Journal of Central Banking* 8. 243-265.

Gabaix, X. (2014): "A Sparcity Based Model of Bounded Rationality", *Quarterly Journal of Economics* 129, 1661-1710.

Gabaix, X. (2017): "Behavioral Macroeconomics via Sparse Dynamic Programming", Working Paper, Harvard.

Gilboa, I. (1987): "Expected Utility with Purely Subjective Non-Additive Probabilities", *Journal of Mathematical Economics* 16, 65-88.

Gilboa, I. and D. Schmeidler (1995): "Case-Based Decision Theory", *Quarterly Journal of Economics* 110, 605-639.

Gul, F. and W. Pesendorfer (2001): "Temptation and Self-Control", *Econometrica*, 69, 1403-143.

Gul, F. and W. Pesendorfer (2004): "Self-Control and the Theory of Consumption", *Econometrica*, 72, 119-158.

Halevy, Y. (2015): "Time consistency: Stationarity and Time Invariance", *Econometrica* 83, 335-352.

Hansen, L.P. and T. Sargent (2001): "Robust Control and Model Uncertainty", *American Economic Review* 91, 60-66.

Hansen, L.P. and T. Sargent (2010): "Fragile Beliefs and the Price of Uncertainty", *Quantitative Economics* 1, 129-162.

Hansen, L.P., T. Sargent and T. Tallarini (1999): "Robust Permanent Income and Pricing", *Review of Economic Studies* 66, 873-907.

Harris, C. and D. Laibson (2001): "Dynamic Choices of Hyperbolic Consumers", *Econometrica* 69, 935-957.

Jensen, M.K. (2007): "Monotone Comparative Statics in Ordered Vector Spaces", *The B.E. Journal of Theoretical Economics* 7 (Topics), Article 35.

Jensen, M.K. (2018): "Distributional Comparative Statics", *Review of Economic Studies* 85, 581-610.

Jensen, M.K. (2018b): "Time Invariant Income-Allocation Problems ", *Mimeo* (available upon request).

Koopmans, T. C. (1960): "Stationary ordinal utility and impatience", *Econometrica* 28, 287-309.

Kreps, D.M. and L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica* 46, 185-200.

Krusell, P. and A. A. Smith (2003), "Consumption–Savings Decisions with Quasi–Geometric Discounting", *Econometrica* 71, 365-375.

Krusell, P., B. Kuruscu, and A. A. Smith (2010), "Temptation and Taxation", *Econometrica* 78, 2063-2084.

Laibson, D. (1996): "Hyperbolic Discount Functions, Undersaving, and Savings Policy", *NBER Working Paper Series* 5635.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting", *Quarterly Journal of Economics* 112, 443-477.

Light, B. (2017): “Uniqueness of equilibrium in a Bewley-Aiyagari model”, Working Paper, Stanford University.

Light, B. and G.Y. Weintraub (2018): “Mean Field Equilibrium: Uniqueness, Existence, and Comparative Statics”, Working Paper, Stanford University.

Loewenstein, G. and D. Prelec (1992): “Anomalies in Intertemporal Choice: Evidence and an Interpretation”, *Quarterly Journal of Economics* 107, 573-598.

Lucas, R.E. and N.E. Stokey (1984): “Optimal Growth with Many Consumers”, *Journal of Economic Theory* 32, 139-171.

Luce, R., *Individual Choice Behavior*, Wiley, New York, 1959.

Machina, M. (1982): “Expected Utility Analysis without the Independence Axiom”, *Econometrica* 50, 277-323.

McFadden, D. (1974): “Conditional logit analysis of qualitative choice behavior”, *in P. Zarembka, ed., Frontiers in Econometrics*, Academic Press, New York, 105-142.

McKelvey, R. And T. Palfrey (1995): “Quantal Response Equilibria for Normal Form Games”, *Games and Economic Behavior* 10, 6-38.

Milgrom, P. (1994): “Comparing Optima: Do Simplifying Assumptions Affect Conclusions?”, *Journal of Political Economy* 102, 607-615.

Milgrom, P. and J. Roberts (1994): “Comparing Equilibria”, *American Economic Review* 84, 441-459.

Milgrom, P. and C. Shannon (1994): “Monotone Comparative Statics”, *Econometrica* 62, 157-180.

M. Nerlove (1958): “Adaptive Expectations and Cobweb Phenomena”, *Quarterly Journal of Economics* 72, 227-240.

Phelps, E. S. and R. A. Pollak (1968): “On Second-Best National Saving and Game Equilibrium Growth”, *Review of Economic Studies* 35, 185â€“199.

Pollak, R.A. (1971): “Additive Utility Functions and Linear Engel Curves”, *Review of Economic Studies* 38, 401-414.

Savage, L.J., *Foundations of Statistics*, Wiley, New York, 1954.

Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity", *Econometrica* 57, 571-587.

Simon, H.A. (1956): "Rational Choice and the Structure of the Environment", *Psychological Review* 63, 129-138.

Sims, C.A. (2003): "Implications of Rational Inattention", *Journal of Monetary Economics* 50, 665-690.

Strotz, R.H. (1956): "Myopia and Inconsistency in Dynamic Utility Maximization", *Review of Economic Studies*, 165-180.

Sun, Y. (2006): "The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks", *Journal of Economic Theory* 126, 31-69.

Tarski, A. (1955): "A Lattice Theoretic Fixed Point Theorem and its Applications", *Pacific Journal of Mathematics* 5, 285-309.

Topkis, D.M. (1978): "Minimizing a Submodular Function on a Lattice", *Operations Research*, 26, 305-321.

Train, K., *Discrete Choice Methods with Simulation*, 2nd Edition, Cambridge University Press, 2009.

Uhlig, H. (1996): "A Law of Large Numbers for Large Economies", *Economic Theory* 8, 41-50.

Vives, X. (1990): "Nash Equilibrium with Strategic Complementarities", *Journal of Mathematical Economics* 19, 305-321.

Woodford, M. (2013): "Macroeconomic Analysis Without the Rational Expectations Hypothesis", *Annual Review of Economics* 5, 303-346.

Quah, J. K.-H. (2007): "The Comparative Statics of Constrained Optimization Problems", *Econometrica* 75, 401-431.

Appendix B: Omitted Proofs (For Online Publication)

Proof of Proposition 3. Consider household $i \in [1-\alpha, 1]$. Under recursive utility with smooth patience function $\rho : c \mapsto R$ (here c is consumption), the mean asset holding a^i given prices w and R is determined as usual from the Euler condition:

$$\rho((1+R)a^i + w) = R \, . \tag{B1}$$

Hence, by the implicit function theorem, a small change in R implies that the asset holding changes to $a^i + da^i$ where:

$$da^i = \frac{(\rho')^{-1} - a^i}{1+R} dR \, . \tag{B2}$$

Now consider a steady state k^* beginning with an arbitrary choice of patience function ρ (of course we must have $(\rho')^{-1} > a^i$ so that a reduction in taxes increases savings, in particular, the household must exhibit increasing marginal impatience in the sense of Epstein and Hynes (1983) and Lucas and Stokey (1984)). Now pick any $\epsilon > 0$ and replace ρ with a new patience function that satisfies (i) $\tilde{\rho}((1+R)a^i + w) = \rho((1+R)a^i + w)$ so that the mean asset holding remains the same (in particular then, the original equilibrium remains the same), and (ii) $\tilde{\rho}'((1+R)a^i + w) = 1/(\epsilon + a^i)$. Inserting into (B2), we see that $da^i = \epsilon/(1+R)$, and so the aggregate change in (mean) assets of households in the set $[1-\alpha, 1]$ is $D(\epsilon) = (1-\alpha)\epsilon/(1+R)$. Since the equilibrium k^* is independent of ϵ, it is clear that for any $\alpha > 0$ we can pick $\epsilon > 0$ such that $D(\epsilon) + C < 0$. Hence for any $\alpha > 0$, we can make the behavioral agents dominate the direct effects given k^* and so turn the tax reduction into a local negative shock. The rest of the proposition follows from our main comparative static results (Theorems 1-4). ■

Proof of Proposition 4. Throughout this proof we omit the superscript θ to simplify notation. We may ignore the monotonic transformation H and write (18) as:

$$s_{w,R,z}(a) \in \arg \max_{a' : (1+R+\sigma)a + wl \geq a'} u_0^{\epsilon}((1+R+\sigma)a + wl - a') + M(a') \, ,$$

where M is a function that does not depend on a. Since $u_0^{\epsilon}((1+R+\sigma)a + wl - a')$ is supermodular in (a, a') if and only if u_0^{ϵ} is concave, it follows from Topkis' theorem (Topkis (1978)) that the smallest and largest optimal savings functions must be increasing in a. Hence the savings correspondence $S_{w,R,z}(a)$ is increasing in the sense of Assumption 1. Next, note that if $S_{w,R,z}(a)$ is upper hemi-continuous in z, it is measurable in z, *i.e.*, the inverse image of every open set is measurable (Aubin and Frankowska (1990), Proposition 8.2.1). To establish both the upper hemi-continuity and measurability requirements of Assumption 1, it therefore suffices to show that under Condition C, $S_{w,R,z}(a)$ is upper hemi-continuous in w, R, a, and z. The proof is the same in each case and in fact, the statement is true if we consider (w, R, a, z) jointly. Nonetheless, to simplify notation we establish the claim only for a. What follows is a sketch only. A detailed proof can be found in Jensen (2018b). We begin with the case where u_0 is strictly concave. Let $a_n \to a$, and $b_n = s_{w,R,z}(a_n) \in S_{w,R,z}(a_n)$ for all n where $s_{w,R,z}$ is a TSSF. Without loss of generality, index again by n a subsequence with $b_n \to b$. We first show that $b \in S_{w,R,z}(a)$. By definition,

$$b_n \in \arg \max_{a':(1+R+\sigma)a_n+wl \geq a'} u_0^\epsilon((1 + R + \sigma)a + wl - a') + M(a')$$

Under Condition C, M is continuous from below (respectively, above) if and only if $s_{w,R,z}(\cdot)$ is continuous from below (respectively, above). In particular, M is continuous if and only if $s_{w,R,z}(\cdot)$ is continuous. For any a, we can pass to yet another subsequence (again indexed by n) such that the convergence $a_n \to a$ is monotone. Since u_0 is strictly concave, $s_{w,R,z}$ is increasing, and it follows then that $b_n \to b$ monotonically. In case (a_n) (and therefore (b_n)) is an increasing sequence, the conclusion that

$$b \in \arg \max_{a':(1+R+\sigma)a+wl \geq a'} u_0^\epsilon((1 + R + \sigma)a + wl - a') + M(a')$$

follows by a standard continuity argument provided that $s_{w,R,z}$, and therefore M is continuous from below. In the second case of decreasing (a_n) and (b_n) the conclusion follows if $s_{w,R,z}$ is continuous from above. Crucially, it may be shown that if $s_{w,R,z}$ is a TSSF, then so is both its lower continuous and its upper continuous closures.[41] Further, since an increasing function is continuous except for at an at most countable number of points, $s_{w,R,z}(a)$ coincides with its lower and upper continuous closures nearly everywhere (as a minimum, at all points

of continuity). Because of this we may from the beginning in the argument above replace $s_{w,R,z}$ with, as appropriate, the lower or upper continuous closure without having to change the sequences (a_n) and (b_n). But then the previous argument may be applied to conclude that $b \in S_{w,R,z}(a)$. If there are only a finite number of TSSFs, this argument implies upper hemi-continuity of $S_{w,R,z}(a)$ (since then for any sequence (a_n) and any sequence (b_n) with $b_n \in S_{w,R,z}(a_n)$ all n, there exist convergent subsequences with $b_n = s_{w,R,z}(a_n)$ all n for some fixed TSSF). If $S_{w,R,z}$ is the union of an infinite family of TSSFs, a slightly more subtle argument is required using that if s^n is a sequence of TSSFs, then its pointwise limit is also a TSSF. Finally, to extend the proof from the case where u_0 is strictly concave to the case where u_0 is merely assumed to be concave, one uses a standard approximation argument: consider a sequence of

[41]Let $s_{w,R,z}$ be an increasing function. Then the lower continuous closure is defined by $\underline{s}_{w,R,z}(a) = \lim_{a^n \uparrow a} s_{w,R,z}(a)$. The upper continuous closure is defined similarly, replacing $a^n \uparrow a$ with $a^n \downarrow a$. These are always well-defined for increasing functions since an increasing function is continuous except at an at most countable number of points (and the points of discontinuity are of the jump type).

strictly concave functions u_0^n that converge pointwise to u_0; repeat the above argument for all n; use that continuity of the maximum operator implies that the limit is optimal for u_0. That $S_{w,R,z}$ has a compact range follows immediately from upper hemi-continuity and boundedness of the set of feasible savings levels. ■

Proof of Proposition 5. We may suppress ϵ and set $\beta = 1$ to simplify notation. Let $V(a,z) = \max_{a'} W(u((1+R+\sigma)a + wl - a'), \mu_{P_z}[V(a',z')])$ denote the value function and note that this is increasing in a under the assumptions of the Proposition. Furthermore V is a continuous function under Condition C. From continuity of the value function, it follows follows that the savings correspondence will be upper hemi-continuous (and hence have a compact range and be measurable in z as explained in the proof of Proposition 4). It remains therefore only to be shown that the savings correspondence is increasing in a. This follows from the argument used in the proof of Proposition 4 if we can show that $W(u((1+R+\sigma)a+wl-a'), \mu_{P_z}[V(a',z')])$ is supermodular in a and a'. Since the objective is concave in a, it is differentiable almost everywhere in a and when the derivative exists it equals: $(1+R+\sigma)W_1'(u((1+R+\sigma)a+wl-a'), \mu_{P_z}[V(a',z')]) \cdot u'((1+R+\sigma)a+wl-a')$. By Theorem 4 in Jensen (2007), it is sufficient for increasing differences/supermodularity in a and a' that this term is increasing in a' between any two points where it is well-defined. Since u is concave, $u'((1+R+\sigma)a+wl-a')$ is increasing in a'. Since $W_1'(u,U)$ is decreasing in u, and increasing in U, and $u((1+R+\sigma)a+wl-a')$ is decreasing in a' and $\mu_{P_z}[V(a',z')]$ is increasing in a', $W_1'(u((1+R+\sigma)a+wl-a'), \mu_{P_z}[V(a',z')])$ is increasing in a'. Since the product of two increasing functions is increasing, the conclusion follows. ■

Proof of Proposition 6. We may take $w = 0$ without loss of generality. In the weakly additive case, utility is

$$u((1-\alpha)(1+R)a) + \beta U(\alpha(1-\alpha_1)(1+R)^2 a, \alpha(1-\alpha_1)\alpha_1(1+R)^3 a, \ldots)$$

Compute the first-order condition and set α_1 equal to α:

$$-(1+R)au'((1-\alpha)(1+R)) + \beta D_a U(\alpha(1-\alpha)(1+R)^2 a, \alpha(1-\alpha)\alpha(1+R)^3 a, \ldots) = 0$$

Note that $D_a U(\alpha(1-\alpha)(1+R)^2 a, \alpha(1-\alpha)\alpha(1+R)^3 a, \ldots) > 0$, so differentiating with respect to β we get something positive. The lhs goes to $+\infty$ as $\alpha \to 0$ and to $-\infty$ as $a \to 1$. Hence the conclusion follows from a standard curve shifting theorem. ■

Proof of Proposition 7. Let $V^{n+1}(a,z;\beta) = \max_{a'} W(u((1+R+\sigma)a+wl-a'), \beta\mu_{P_z}[V^n(a',z';\beta)])$ and note that if V^n is increasing in β, then V^{n+1} is increasing in β. Further, $D_a V^{n+1}(a,z,\beta) =$

$D_1 W(u((1+R+\sigma)a+wl-a'), \beta\mu_{P_z}[V^n(a', z'; \beta)]) \cdot (1+R+\sigma)$, and since $W(u, U)$ is supermodular in u and U, $D_a V^{n+1}(a, z, \beta)$ is increasing in β if and only if $\beta\mu_{P_z}[V^n(a', z'; \beta)]$ is increasing in β. A sufficient condition for $\beta\mu_{P_z}[V^n(a', z'; \beta)]$ to be increasing in β is that $V^n(a', z'; \beta)$ is increasing in β for all a' and z'. By iteration, we conclude that the value function $V(a, z; \beta)$ is supermodular in a and β. From this the conclusion follows by the same argument as that used in the proof of Proposition 5. ■