NBER WORKING PAPER SERIES

A THEORY OF POLITICAL AND ECONOMIC CYCLES

Laurence Ales
Pricila Maziero
Pierre Yared

Working Paper 18354
http://www.nber.org/papers/w18354

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2012

We would like to thank Marina Azzimonti, Sandeep Baliga, Chris Blattman, V.V. Chari, Mike Golosov, Brett House, Larry Jones, Emi Nakamura, Torsten Persson, Aleh Tsyvinski and seminar participants at Stanford, the Nemmers Prize Conference on The Political Economy of Growth and Development, the Workshop on Macroeconomic Applications of Dynamic Games and Contracts and the SED for comments. We would like to thank Sergei Kolbin for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

A Theory of Political and Economic Cycles
Laurence Ales, Pricila Maziero, and Pierre Yared
NBER Working Paper No. 18354
September 2012
JEL No. D82,E62,H21,P16

ABSTRACT

We develop a theoretical framework in which political and economic cycles are jointly determined. These cycles are driven by three political economy frictions: policymakers are non-benevolent, they cannot commit to policies, and they have private information about the tightness of the government budget and rents. Our first main result is that, in the best sustainable equilibrium, distortions to production emerge and never disappear even in the long run. This result is driven by the interaction of limited commitment and private information on the side of the policymaker, since in the absence of either friction, there are no long run distortions to production. Our second result is that, if the variance of private information is sufficiently large, there is equilibrium turnover in the long run so that political cycles never disappear. Finally, our model produces a long run distribution of taxes, distortions, and turnover, where these all respond persistently to temporary economic shocks. We show that the model's predictions are consistent with the empirical evidence on the interaction of political and economic cycles in developing countries.

Laurence Ales
Tepper School of Business
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
ales@cmu.edu

Pricila Maziero
University of Pennsylvania
3620 Locust Walk SHDH 2324
Philadelphia, PA, 19104
maziero@wharton.upenn.edu

Pierre Yared
Columbia University
Graduate School of Business
3022 Broadway, Uris Hall 823
New York, NY 10027
and NBER
pyared@columbia.edu

1 Introduction

The collapse of commodity prices in the late 1970s and early 1980s caused a sharp decline in government revenues in many sub-Saharan African countries. Unable to fund public services, leaders faced the threat of removal. In some cases, they responded to this threat by taking measures which increased social programs while simultaneously expropriating private enterprises, further exacerbating the economic crisis.[1]

This episode is not unique but more generally reflects the fact that economic and political cycles are deeply interconnected. On the one hand, economic shocks impact the tenure of leaders, as incumbents are often replaced following negative economic shocks. On the other hand, political risk and the threat of turnover can often induce policymakers facing potential replacement to become shortsighted and to choose inefficient policies.[2]

In this paper, we develop a framework in which political and economic cycles are jointly determined. In our environment, these cycles are driven by three key political economy frictions. First, policymakers are not benevolent, and are instead driven by political rents and by the desire to preserve power. Second, policymakers lack commitment, and once in office, they are not bound to the promises which they made to citizens. Finally, policymakers have private information about the tightness of the government budget and their rent-seeking activities. We embed these frictions in an environment which combines two frameworks. The first framework is a standard political accountability model with asymmetric information in which citizens can punish incumbents with replacement. The second framework is a dynamic production economy with rent-seeking.

More formally, our economy is populated by households which choose investment and a non-benevolent policymaker who chooses taxes and rents. The policymaker cannot commit to policies after households have made their investment decision, and households discipline the policymaker by threatening to replace him. The government controls a stochastic endowment, where this captures a shock to the value of government royalties or to the cost of public spending. The policymaker privately observes the size of this shock and privately chooses the level of rents. This implies that if citizens observe high taxes, they may not be able to determine whether this is due to an exogenous aggregate shock which tightened the budget or whether this is due to unobserved rent-seeking by the policymaker.

We consider the best sustainable equilibrium which maximizes the ex-ante welfare of citizens. This equilibrium takes into account the joint interaction of the constraints of limited

[1] See Bates (2008) for further discussion of these episodes. As an example, following the collapse of copper prices, President Kaunda of Zambia nationalized several milling companies, imposed price controls, and limited government debt service as part of the Interim New Economic Recovery Programme. Between 1988 and 1991, investment in Zambia declined by 17%. See Baylies and Szeftel (1992) and Simutanyi (1996) for additional discussion.

[2] As an another example, many Latin American countries dependent on commodity exports experienced economic and political crises following the collapse of commodity prices in the late 1970s and early 1980s. For a discussion of the experience of Mexico, see Bergoeing, Kehoe, Kehoe, and Soto (2002), for example.

commitment and private information on the side of the policymaker. We show how in the absence of either friction, there are no distortions to production since the level of investment is efficiently chosen in the long run. In the absence of asymmetric information, for instance, our model features *backloading*. Specifically, a policymaker is never replaced, though if he deviates by expropriating households, he is replaced off the equilibrium path. While distortions emerge along the equilibrium path in order to limit the resources which can be expropriated by the policymaker, these distortions eventually disappear as rents rise and reduce his incentives to expropriate. Note that the absence of long run distortions under full information not unique to our model, but common across a large class of full information principal-agent environments in which the agent suffers from limited commitment, as in Acemoglu, Golosov, and Tsyvinski (2008, 2010a,b), for example.[3] Analogously, under asymmetric information and in the presence of full commitment, there are never distortions to production. Because the policymaker has limited discretion over the choice of taxes under full commitment, the payoffs from his decisions are independent of the level investment. As such, distortions to production cannot facilitate incentive provision and they never appear. Therefore, under either full information or full commitment, there are no long run distortions to production.

The first main result of our paper is that distortions to production emerge and *never* disappear, even in the long run. This feature of our model is a consequence of the joint interaction of the limited commitment and the asymmetric information frictions. This result is due to the fact that a policymaker is always provided with dynamic incentives to not privately rent-seek. More specifically, if a shock tightens (slackens) the budget constraint so that observed taxes are high (low), then the policymaker is punished (rewarded) in the future with lower (higher) payment. Eventually a long sequence of negative shocks push payments to the policymaker sufficiently down that the policymaker becomes tempted to fully expropriate the investment of households. Anticipating this threat, households invest less, so that distortions to production eventually emerge as a means of preventing full expropriation. This result arises as a consequence of optimality and *not* feasibility since allocations in which there are no distortions to production are sustainable in our environment; however, they are suboptimal since they do not entail enough risk-sharing between households and the policymaker. Importantly, this result holds for any variance in the private information of the policymaker. Therefore, the introduction any arbitrarily small amount of privately observed uncertainty to the full information benchmark leads to the presence of long run distortions, completely altering the predictions of the full information benchmark.

The second main result of our paper is that there is turnover in the long run if the variance of the private information of the policymaker is sufficiently large. This is because, if the variance of private information is large, then the policymaker has high private rent-seeking opportunities, and replacement is a useful means of preventing private rent-seeking. More

[3] This is conditional on both the principal and the agent having the same discount factor. For other examples, see also Thomas and Worrall (1994), Ray (2002) and Albuquerque and Hopenhayn (2004), among others.

specifically, society has two tools for providing incentives to policymakers to not privately rent-seek. On the one hand, society can directly pay higher future rents to reward policymakers who chooses low taxes today. Though this costs societal resources, it reduces the policymaker's incentives to fully expropriate households since he values preserving power, and it allows households to choose the efficient level of investment today. On the other hand, society can instead punish policymakers who choose high taxes by removing them from office in the future. This does not cost any societal resources, but it raises a policymaker's incentives to fully expropriate households today since the horizon of the policymaker is reduced. In response, households are forced to invest less today, causing economic distortions. If the variance of private information is large, then a policymaker has high private rent-seeking opportunities, and providing incentives to the policymaker via payments alone is extremely costly. In this situation, the use of replacement is efficient–despite its effect on increasing economic distortions–as it allows society to make smaller payments to the policymaker. This result effectively generalizes the endogenous turnover result of Ferejohn (1986) to an economy in which production is determined by optimizing households and where policymakers and citizens choose fully history dependent strategies associated with the best sustainable equilibrium.[4]

The final result of our paper is that our model generates a long run distribution of taxes, distortions, and turnover. In particular, we show that negative (positive) economic shocks which tighten (slacken) the government budget lead to a reduction (increase) in future taxes, investment, and tenure, where these all respond persistently to temporary economic shocks. Moreover, the model predicts that periods of possible turnover are associated with the lowest equilibrium taxes and the highest equilibrium investment distortions. Finally, these dynamics are associated with a probability of turnover which is a negative function of the tenure length of the incumbent. Note that these long run dynamics are significantly different relative to those in an environment with full information, since in such an environment, taxes are i.i.d., there are no distortions, and there is no turnover in the long run.

Though the main focus of this paper is theoretical, many of the predictions of the model are consistent with existing qualitative empirical patterns in previous research. We supplement this previous work by considering the effect of commodity price shocks in developing countries. As has been well documented, commodity price shocks are an important source of business cycles in developing countries (e.g., Deaton (1999)). In terms of our model, commodities are a funding source for many governments in developing countries, so that global shocks to commodity prices outside of policymakers' control can tighten or slacken the government budget constraint. We find a number of empirical regularities which are consistent with the predictions of our model. Specifically, commodity price shocks which clearly affect

[4] Ferejohn (1986) considers an environment in which a policymaker can only be punished or rewarded with replacement and in which citizens choose Markovian strategies. The presence of turnover in his environment does not require a sufficiently large variance in the private information of the policymaker, and this is because the model does not allow for production or distortions.

income and investment also reduce the tenure of leaders and do so in a persistent fashion. Moreover, periods of turnover are associated with significantly depressed investment growth and tax revenue growth. Finally, the probability of turnover in the data is decreasing in the tenure length of the incumbent.

Our paper is connected to a very large literature which studies the effect of political uncertainty on fiscal policy distortions. In this literature, the presence of political uncertainty leads policymakers to be short-sighted and thus choose inefficient policies.[5] Our main contribution to this literature is that we endogenize the level of political uncertainty so that it is time-varying and a function of the entire history of economic shocks. As such, our study provides a framework in which the level of distortions and the level of turnover are jointly determined endogenously. In this regard, our paper is very closely related to the literature on the political business cycle, and in particular to the work of Rogoff (1990).[6] He endogenizes political uncertainty in a three-period economy in which office-driven policymakers have private information about their competency, so that voting is prospective. In contrast to this work, we consider a setting in which policymakers are identical but have private information about the temporary state of the economy and their rent-seeking activities, so that voting is retrospective. This facilitates characterization of the best sustainable equilibrium in a fully dynamic infinite horizon economy.[7] Our paper is also related to the literature on retrospective voting, going back to the seminal work of Barro (1973) and Ferejohn (1986).[8] We contribute to this literature by characterizing the dynamics of turnover in a dynamic production economy with optimizing households in which citizens choose history-dependent non-Markovian strategies. Finally, our analysis contributes to a large literature on dynamic contracts. Our main contribution to this literature is to explore the long run implications of the interaction of a model of limited commitment which features backloading and a model of asymmetric information.[9] We show that while neither framework on its own leads to long run distortions,

[5]This theme emerges in a large body of work, which includes, but is by no means only limited to Persson and Svensson (1989), Alesina and Tabellini (1990), Alesina and Perotti (1994), Krusell and Rios-Rull (1999), Battaglini and Coate (2008), Song, Storesletten, and Zilibotti (2009), Azzimonti (2010),Caballero and Yared (2010) and Aguiar and Amador (2011).

[6]See Alesina (1988) and Alesina, Roubini, and Cohen (1997) as well as Drazen (2000) for an overview of the political business cycle literature together with relevant references. In contrast to the majority of this work, we focus on the fiscal as opposed to the monetary channel for political distortions. Additionally, we consider a fully dynamic economy with retrospective voting so that turnover risk is completely endogenous and not exogenously determined by the timing of elections. For this reason, in our discussion of the empirical evidence, we focus on developing economies where turnover is not consistently determined by regularly held elections.

[7]As discussed in Rogoff (1990), it is very difficult to analyze a prospective voting framework in a fully dynamic environment.

[8]See also Banks and Sundaram (1998), Persson and Tabellini (2000), Besley (2006), Egorov (2009), and Fearon (2010) for extensions.

[9]For some examples of full information models which feature backloading, see the work cited in Footnote 3. For models which feature asymmetric information, see Thomas and Worrall (1990) and Atkeson and Lucas (1992), among others. Li and Matouschek (2011) consider a related environment with limited commitment and asymmetric information and find the presence of long run distortion. However, in contrast to our work, their result is not driven by optimality considerations but by the non-existence of *any* equilibrium without long run distortions.

the interaction of the two frameworks does lead to long run distortions.[10]

This paper is organized as follows. Section 2 describes the model. Section 3 defines and provides a recursive representation for the equilibrium. Section 4 characterizes the benchmark cases with full information and full commitment. Section 5 summarizes our results once the frictions of limited commitment and asymmetric information are allowed to interact. Section 6 discusses empirical evidence consistent with the predictions of the model. Section 7 concludes. The Appendix includes proofs and additional material not included in the text.

2 Model

We describe an environment in which households choose a level of investment and policies are chosen by self-interested policymakers. Policymakers cannot commit to policies, have private information about the shocks to the government budget, and can privately rent-seek. In this environment, households discipline policymakers by threatening to remove them from power.

2.1 Economic Environment

There are discrete time periods $t = \{0, ..., \infty\}$. In every period there is a stochastic state $\theta_t \in \Theta \equiv \{\theta^1, ..., \theta^N\}$ with $\theta^n > \theta^{n-1} \geq 0$ and $N \geq 2$. The state is i.i.d. and occurs with probability $\pi(\theta_t)$. There is a continuum of mass 1 of identical households with the following utility:

$$E_0 \left(\sum_{t=0}^{\infty} \beta^t u(c_t) \right), \; \beta \in (0,1), \tag{1}$$

where c_t is consumption and β is the discount factor. $u(\cdot)$ is strictly increasing and strictly concave in c_t with $\lim_{c \to 0} u'(\cdot) = \infty$ and $\lim_{c \to \infty} u'(\cdot) = 0$. Households enter every period with a fixed endowment $\omega > 0$. They decide how much of this endowment to dedicate to investment $i_t \geq 0$ which produces output $y_t = f(i_t)$. $f(\cdot)$ is strictly increasing and strictly concave in i_t with $f(0) = 0$, $\lim_{i \to 0} f'(\cdot) = \infty$ and $\lim_{i \to \infty} f'(\cdot) = 0$. A household has the following per period budget constraint:

$$c_t = \omega - i_t + y_t - \tau_t \left(y^t\right) \quad \forall t, \tag{2}$$

where $\tau_t \left(y^t\right) \gtreqless 0$ represents the taxes incurred which can be a function of the entire history of output by the household y^t. We constrain taxes so that $\tau_t \left(y^t\right) \leq y_t$, meaning that the government cannot impose a tax on production which exceeds one hundred percent. Note

[10] Our paper is also related to several lines of research which consider the role of private government information (e.g. Sleet (2001, 2004), Athey, Atkeson, and Kehoe (2005) and Amador, Werning, and Angeletos (2006)). The main departure from this work is our focus on an environment with a non-benevolent government in which citizens can punish the policymaker with replacement.

that independently of the level of taxes, a household can always guarantee itself a level of consumption of at least ω by choosing investment to equal 0.

There is a continuum of potential and identical self-interested policymakers each indexed by $j \in J$. Let $P_{jt} = \{0, 1\}$ be an indicator function which denotes whether a policymaker j has power in period t where $P_{jt} = 1$ denotes that policymaker j holds power. Only one policymaker holds power, so that if $P_{jt} = 1$ then $P_{-jt} = 0$ for $-j \neq j$. Policymaker j has the following utility:

$$E_0 \left(\sum_{t=0}^{\infty} \beta^t \left(P_{jt} v\left(x_t\right) + \left(1 - P_{jt}\right) \underline{V} \left(1 - \beta\right)\right) \right), \tag{3}$$

for $x_t \geq 0$ which represents rents paid to the policymaker in power and $\underline{V}(1-\beta) \leq v(0)$ which represents the exogenous flow utility to a policymaker who is not in power. $v(\cdot)$ is strictly increasing and strictly concave in x_t with $\lim_{x \to 0} v'(\cdot) = \infty$ and $\lim_{x \to \infty} v'(\cdot) = 0$.

The government has the following per period budget constraint:

$$x_t = \tau_t \left(y^t\right) + \theta_t, \tag{4}$$

where we have taken into account that since households are identical, the government's aggregate tax revenue equals the individual tax burden $\tau_t\left(y^t\right)$. θ_t represents a government endowment which is determined after investment is undertaken and before policies $\tau_t\left(y^t\right)$ are chosen. It captures a shock to the cost of public spending or to the value of government royalties.[11] The resource constraint of the economy implied by (2) and (4) is:

$$c_t + x_t = \omega - i_t + y_t + \theta_t. \tag{5}$$

The most important feature of this setting is that while the entire society observes the policy $\tau_t\left(y^t\right)$, the values of x_t and θ_t are *privately* observed by the policymaker in power. This means that citizens cannot distinguish between resources which are used to alleviate the government budget constraint from resources which are used for private rent-seeking by the policymaker.

2.2 Political Environment

The political environment is as follows. At every date t, citizens decide whether or not to replace an incumbent. Formally, if $P_{jt-1} = 1$, then if citizens choose $P_{jt} = 1$ policymaker j remains in power, and if citizens choose $P_{jt} = 0$ a replacement policymaker $k \in J$ is randomly chosen to replace j from the set J (i.e., nature stochastically chooses $P_{kt} = 1$ for some $k \in J$).

[11] We do not allow θ_t to be negative as a technicality in order to guarantee that $x_t \geq 0$ for any level of investment. If we allow $f(0)$ to be positive, then θ_t can be as low as $-f(0)$. Alternatively, if θ_t is always negative so that it is interpreted as a public good and $f(0) = 0$, then it is possible to allow the government to not provide any public goods at an arbitrarily large cost to the households, so that this only happens off the equilibrium path. This extension does not alter any of our results. Details available upon request.

To reduce notation, we let $P_t = \{0, 1\}$ correspond to the decision of whether or not keep an incumbent at date t.[12]

Following the replacement decision, households make their investment i_t. Nature then draws θ_t which is privately observed by the policymaker. The policymaker then chooses policies $\{x_t, \tau_t(y^t)\}$ subject to (4) and subject to the constraint that $\tau_t(y^t) \leq y_t$. Note that a policymaker can always choose $\tau_t(y^t) = y_t$ after the household investment decision has been determined, implying from (5) that $c_t = \omega - i_t$. Note that this value may be negative, and in this circumstance, we define $u(c_t) = -\infty$.[13]

A key feature of this game is that even though citizens make their economic decisions independently, they make their political decisions regarding the replacement of the policymaker jointly. Since citizens are identical, there is no conflict of interest between them. These joint political decisions can be achieved by a variety of formal or informal procedures such as elections, protests, revolutions, or coups. We simplify the discussion by assuming that the decision is taken by the same single representative citizen in every period.[14]

There are two essential features of this game. First, the policymaker suffers from limited commitment within the period. Specifically, following the investment decision of households, the policymaker may decide to fully expropriate households and set rents equal to $y_t + \theta_t$, which is the maximum. Second, the policymaker privately observes the government budget shock and the total amount of rent-seeking. As such, if the shock θ_t is high so that the government budget is slack and taxes can be low, the policymaker may instead pretend that the government budget is tight so as to choose higher taxes and to privately rent-seek. In the following section, we investigate how reputational considerations can alleviate the problem of limited commitment and asymmetric information in this environment.

3 Best Sustainable Equilibrium

As in Chari and Kehoe (1993a,b) we consider sustainable equilibria. Individual households are anonymous and non-strategic in their private market behavior, though the representative citizen is strategic in his replacement decision. The policymaker in power is strategic in his choice of policies, and he must ensure that the government's budget constraint is satisfied given the resource constraint and the anonymous market behavior of households. Using this definition, we characterize the entire set of sustainable equilibria and we consider conditions which are necessary in the efficient sustainable equilibrium.

[12] In our model, policymakers can only be in power once. Nonetheless, one can extend our analysis under some refinements so as to allow for the possibility of returning to power without altering any of our main results. Under this extension, $\underline{V}$ represents an endogenous value of being thrown out of power. Details available upon request.

[13] Though negative household consumption will never occur along the equilibrium path, it could in principle occur off the equilibrium path if the policymaker decides to fully expropriate households. Note that our main results can be generalized to an environment in which the household's utility function is well defined for any arbitrarily negative level of consumption.

[14] This is identical to the decision being made via majoritarian elections with sincere voting.

3.1 Definition of Sustainable Equilibrium

We begin by defining strategies of the citizens and the policymaker. We introduce a publicly observed random variable to allow for correlated strategies. In every period, $z_t \in Z \equiv [0,1]$ is drawn from a uniform distribution. This publicly observed random variable allows citizens to probabilistically replace an incumbent.

Define $h_t^0 = \{z^t, \{P_j^{t-1}\}_{j\in J}, \rho^{t-1}\}$ as the history of the public random variable, replacement decisions, and policies after the realization of z_t, where ρ_t corresponds to the vector of tax policies for each y^t at date t. Let $h_t^1 = \{h_t^0, \{P_j^t\}_{j\in J}\}$ and let $h_t^2 = \{h_t^1, \{P_j^t\}_{j\in J}, \theta_t\}$, where h_t^2 is only observed by the incumbent policymaker. A representative citizen's replacement strategy Υ assigns a replacement decision for every h_t^0. A representative household's investment sequence Φ assigns a level of investment at every h_t^1. The incumbent policymaker's strategy Σ assigns policies for every h_t^2. Let $\Upsilon|_{h_t^0}$ represent the continuation strategy of the representative citizen at h_t^0 and define $\Phi|_{h_t^1}$ and $\Sigma|_{h_t^2}$ analogously.[15]

The representative citizen's replacement strategy Υ solves the representative citizen's problem if, at every h_t^0, the continuation strategy $\Upsilon|_{h_t^0}$ maximizes household welfare given $\{\Phi, \Sigma\}$. A representative household's investment sequence Φ solves the representative household's problem if at every h_t^1, the continuation investment sequence $\Phi|_{h_t^1}$ maximizes household welfare given $\{\Upsilon, \Sigma\}$ and given the household's budget constraint. The incumbent policymaker's strategy Σ solves the incumbent policymaker's problem if, at every h_t^2, the continuation strategy $\Sigma|_{h_t^2}$ maximizes the incumbent policymaker's welfare given $\{\Upsilon, \Phi\}$ and given the government's budget constraint and the maximum constraint on taxes. Note that because households are anonymous, public decisions are not conditioned on their allocation.

A sustainable equilibrium consists of $\{\Upsilon, \Phi, \Sigma\}$ for which Υ solves the representative citizen's problem, Φ solves the household's problem, and Σ solves the incumbent policymaker's problem.

3.2 Sustainable Equilibrium Allocations

To characterize the best sustainable equilibrium, we first characterize the set of sustainable allocations supported by sustainable equilibrium strategies. Let $q_t = \{z_0, ..., z_{t-1}, \theta_0, ..., \theta_{t-1}\}$, the *exogenous* equilibrium history of public signals and states prior to the realization of z_t. With some abuse of notation, define an equilibrium allocation as a function of the exogenous history:

$$\delta = \{P_t(q_t, z_t), i_t(q_t, z_t), c_t(q_t, z_t, \theta_t), x_t(q_t, z_t, \theta_t)\}_{t=0}^{\infty}, \tag{6}$$

[15] We are implicitly assuming that households choose identical investment strategies and that policymakers also choose identical strategies independently of their identity. These assumptions are without loss of generality since we focus on the best sustainable equilibrium for households.

where $P_t(q_t, z_t)$ is the value of P_t chosen at q_t, z_t and the other variables are defined analogously. Define

$$V_t(q_t) = \int_0^1 \left[\begin{array}{c} (1 - P_t(q_t, z_t))\,\underline{V} + \\ P_t(q_t, z_t)\left(\sum_{\theta_t \in \Theta} \pi(\theta_t)\left(v\left(x_t(q_t, z_t, \theta_t)\right) + \beta V_{t+1}(q_t, z_t, \theta_t)\right)\right) \end{array} \right] dz_t,$$

the welfare expected by the incumbent at the beginning of the stage game prior to the realization of the public signal z_t. Moreover, define $J_t(q_t)$ analogously as the welfare of the households prior to the realization of z_t:

$$J_t(q_t) = \int_0^1 \left[\sum_{\theta_t \in \Theta} \pi(\theta_t)\left(u\left(c_t(q_t, z_t, \theta_t)\right) + \beta J_{t+1}(q_t, z_t, \theta_t)\right) \right] dz_t.$$

Finally, let $\mathcal{F}$ be the set of feasible allocations defined as follows. $\delta \in \mathcal{F}$ if and only if every element of δ at $\{q_t, z_t\}$ is measurable with respect to public information up to t and for all $\{q_t, z_t, \theta_t\}$, δ satisfies the following constraints:

$$P_t(q_t, z_t) \in \{0, 1\},\ i_t(q_t, z_t) \geq 0,\ c_t(q_t, z_t, \theta_t) \geq 0,\ x_t(q_t, z_t, \theta_t) \geq 0,$$

$$c_t(q_t, z_t, \theta_t) + x_t(q_t, z_t, \theta_t) = \omega - i_t(q_t, z_t) + f\left(i_t(q_t, z_t)\right) + \theta_t, \text{ and} \tag{7}$$

$$x_t(q_t, z_t, \theta_t) \leq f\left(i_t(q_t, z_t)\right) + \theta_t. \tag{8}$$

The following proposition provides necessary and sufficient conditions for an allocation to be supported by sustainable equilibrium strategies.

Proposition 1 ***(sustainable equilibrium allocation)*** *δ is supported by sustainable equilibrium strategies if and only if $\delta \in \mathcal{F}$ and $\forall q_t, z_t$*

$$v\left(x_t(q_t, z_t, \theta_t)\right) + \beta V_{t+1}(q_t, z_t, \theta_t) \geq v(x_t(q_t, z_t, \widehat{\theta}) + \theta_t - \widehat{\theta}) + \beta V_{t+1}(q_t, z_t, \widehat{\theta}) \forall\, \theta_t, \widehat{\theta} \in \Theta, \tag{9}$$

$$v\left(x_t(q_t, z_t, \theta_t)\right) + \beta V_{t+1}(q_t, z_t, \theta_t) \geq v\left(f\left(i_t(q_t, z_t)\right) + \theta_t\right) + \beta \underline{V} \quad \forall\, \theta_t \in \Theta, \text{ and} \tag{10}$$

$$\sum_{\theta_t \in \Theta} \pi(\theta_t)\left(u\left(c_t(q_t, z_t, \theta_t)\right) + \beta J_{t+1}(q_t, z_t, \theta_t)\right) \geq u(\omega) / (1 - \beta). \tag{11}$$

Proof. See Appendix. ■

The intuition for Proposition 1 is as follows. The government has significant flexibility in choosing its non-linear tax instrument $\tau_t(y^t)$. This effectively implies that as long as an allocation satisfies $\delta \in \mathcal{F}$ and (11), there exists a tax policy which implements the allocation. Intuitively, the government can effectively induce households to invest any amount as long as their expected consumption under the policy weakly exceeds that under 0 investment forever which yield $u(\omega) / (1 - \beta)$. This explains why the constraint that $\delta \in \mathcal{F}$ and that (11) is satisfied is necessary and sufficient to guarantee optimality on the side of the households.[16]

Constraints (9) and (10) capture the incentive compatibility constraints on the side of the policymaker. More specifically, constraint (9) captures the private information of the government. It guarantees that, if the policymaker is prescribed a particular policy given the realized shock θ_t, he does not instead privately choose an alternative policy appropriate for another shock $\widehat{\theta}$ which has not been realized. Given (4), such an alternative policy provides him with rents equal to $x_t(q_t, z_t, \widehat{\theta}) + \theta_t - \widehat{\theta}$ at t and a continuation value of $V_{t+1}(q_t, z_t, \widehat{\theta})$ at $t+1$. Constraint (9) guarantees that he weakly prefers to choose the prescribed policy which provides him with rents equal to $x_t(q_t, z_t, \theta_t)$ at t and a continuation value of $V_{t+1}(q_t, z_t, \theta_t)$ at $t+1$. Constraint (10) captures the additional constraint of limited commitment. At any date t, the policymaker can engage in an observable deviation by expropriating all of the output of the economy. In this situation, this constraint guarantees that he prefers to pursue prescribed policies versus making this observable deviation and being thrown out of power which provides him with welfare $\underline{V}$ from tomorrow onward.[17]

A natural question emerges regarding the citizens' incentives to follow the prescribed replacement rules. Proposition 1 shows that satisfaction of such incentives does not place restrictions on the set of sustainable allocations δ. This is because policymakers are identical, which means that citizens can always be made indifferent on the margin between the current incumbent and any replacement policymaker.[18]

Let Λ represent the set of allocations $\delta \in \mathcal{F}$ which satisfy conditions $(9) - (11)$. The best sustainable equilibrium in our environment is a solution to the following program:

$$\max_{\delta \in \Lambda} E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(q_t, z_t, \theta_t\right)\right), \tag{12}$$

where the additional constraint that $\delta \in \Lambda$ ensures that the allocation satisfies sustainability constraints. Note that this definition is analogous to that of Acemoglu, Golosov, and Tsyvinski (2008, 2010a,b) since it ignores the welfare of the incumbent as well as all candidate policymakers.

[16]Note that if taxes could not be history dependent and could only depend on y_t, then (11) would be replaced by $\sum_{\theta_t \in \Theta} \pi(\theta_t) u(c_t(q_t, z_t, \theta_t)) \geq u(\omega) / (1-\beta)$. The analysis under this modified constraint is complicated by the fact that the implied value function is no longer necessarily differentiable. In the cases where it is differentiable, all of our results are preserved. Details available upon request.

[17]As a reminder, $\underline{V} \leq v(0) / (1-\beta)$ so that there is no worse punishment than being thrown out of office.

[18]In equilibrium, households could also strictly prefer to pursue the prescribed replacement rules if future policymakers punish households for deviating from these rules with full expropriation in the future. What is critical here is that candidate policymakers observe the history of the game and can therefore determine if citizens deviated from the equilibrium replacement rule. Note that, by this rationale, our main results are also preserved if we allow for an exogenous cost of replacing incumbents, as long as this cost is sufficiently small. Details available upon request.

3.3 Recursive Representation of Best Sustainable Equilibrium

To facilitate the analysis, we provide a recursive formulation for (12). Define $\overline{J}$ as the utility attained under the solution to (12). Note that if the solution to (12) admits $P_t(q_t, z_t) = 0$ for some $\{q_t, z_t\}$, then the welfare of households at $\{q_t, z_t\}$ is equal to $\overline{J}$. This is because if it were not the case, it would be possible to pursue the same sequence of allocations from $\{q_t, z_t\}$ onward as those starting from date 0, and this would continue to satisfy all of the sustainability constraints while strictly increasing the welfare of households. Therefore, whenever a policymaker is replaced, households receive their highest continuation value $\overline{J}$.

A natural question pertains to the continuation value that a policymaker receives in his first period in power. In principle, it is possible that (12) admits different levels of welfare for new incumbents even though households continue to receive $\overline{J}$. In this situation, we select the equilibrium which also maximizes the welfare of the policymaker subject to providing the households with their maximum welfare $\overline{J}$, where we denote this welfare by V_0. Therefore, the equilibrium resets whenever turnover occurs.[19]

Let $J(V)$ correspond to the highest continuation value which the households receive at t conditional on having promised the $t-1$ policymaker a continuation value V starting from date t. Starting from a given V, let α correspond to

$$\alpha = \left\{ P(z) \in \{0,1\}, i(z) \geq 0, \left\{ c(\theta,z) \geq 0, x(\theta,z) \geq 0, V'(\theta,z) \right\}_{\theta \in \Theta} \right\}_{z \in [0,1]}, \tag{13}$$

where $P(z)$ is value of P_t chosen if $z_t = z$, and $i(z)$, $c(\theta,z)$, and $x(\theta,z)$ are analogously defined. Let $V'(\theta,z)$ correspond to the continuation value starting from $t+1$ if $z_t = z$ and $\theta_t = \theta$. Moreover, let $\overline{V}$ correspond to the highest continuation value which can be provided

[19]This is consistent with the notion of constrained Pareto efficiency which we are using. In practice, the cases we consider will imply a unique V_0, so that this multiplicity is not an issue for any of the results in our paper.

to the incumbent policymaker in a sustainable equilibrium. The recursive program is:

$$J(V)=\max_{\alpha}\left\{\int_0^1\left[\begin{array}{c}(1-P(z))\,\overline{J}+\\ P(z)\left(\sum_{\theta\in\Theta}\pi(\theta)\left(u(c(\theta,z))+\beta J(V'(\theta,z))\right)\right)\end{array}\right]dz\right\} \tag{14}$$

s.t.

$$V=\int_0^1\left[(1-P(z))\,\underline{V}+P(z)\left(\sum_{\theta\in\Theta}\pi(\theta)\left(v(x(\theta,z))+\beta V'(\theta,z)\right)\right)\right]dz, \tag{15}$$

$$c(\theta,z)+x(\theta,z)=\omega-i(z)+f(i(z))+\theta\ \forall\theta,z, \tag{16}$$

$$x(\theta,z)\leq f(i(z))+\theta\ \forall\theta,z, \tag{17}$$

$$v(x(\theta,z))+\beta V'(\theta,z)\geq v(x(\widehat{\theta},z)+\theta-\widehat{\theta})+\beta V'(\widehat{\theta},z)\ \forall\theta,\widehat{\theta},z, \tag{18}$$

$$v(x(\theta,z))+\beta V'(\theta,z)\geq v(f(i(z))+\theta)+\beta\underline{V}\ \forall\theta,z, \tag{19}$$

$$\sum_{\theta\in\Theta}\pi(\theta)\,u\left(c(\theta,z)+\beta J\left(V'(\theta,z)\right)\right)\geq u(\omega)/(1-\beta)\ \ \forall z, \tag{20}$$

$$\text{and } V'(\theta,z)\in\left[\underline{V},\overline{V}\right]\ \ \forall\theta,z. \tag{21}$$

(14) takes into account that if $P(z)=0$, the incumbent policymaker is replaced and households receive a continuation welfare $\overline{J}$. Otherwise, the incumbent is not replaced and the households receive consumption $c(\theta,z)$ today and a continuation value $J(V'(\theta,z))$ starting from tomorrow for each θ,z. Constraint (15) is the promise keeping constraint for the current incumbent which guarantees that his continuation value equals V. It takes into account that if he is replaced, he receives a continuation value $\underline{V}$. If he is not replaced, he receives consumption $x(\theta,z)$ today and a continuation value $V'(\theta,z)$ starting from tomorrow for each θ,z. Constraints (16) − (20) correspond to the recursive versions of constraints (7) − (11). Constraint (21) guarantees that the continuation value $V'(\theta,z)$ is in the feasible range between $\underline{V}$ and $\overline{V}$.[20]

In the Appendix, we establish that $J(V)$ is weakly concave in V and that it is continuously differentiable in $V\in\left(\underline{V},\overline{V}\right)$. In addition, it has the following property: It satisfies $J(V)=\overline{J}$ for $V\in[\underline{V},V_0]$ and it is strictly decreasing in V if $V\in\left(V_0,\overline{V}\right]$. That $J(V)$ is weakly decreasing follows from the fact that it must not be possible to make households strictly better off without making the incumbent weakly worse off, and this follows from the definition of the best sustainable equilibrium. If $V\in(\underline{V},V_0)$, then the incumbent policymaker faces a positive probability of replacement, and in this situation households randomize between keeping the policymaker in power which provides him with V_0 or throwing the policymaker out of power which provides him with $\underline{V}$. In both of these circumstances, households receive a continuation welfare equal to $\overline{J}$ and the policymaker who is ultimately in power–whether it is last period's incumbent or a replacement policymaker–receives a continuation values of V_0 (conditional on

[20] Note that in addition, it must be the case that if $c(\theta',z)=c(\theta'',z)$ for $\theta'\neq\theta''$, then $V'(\theta',z)=V'(\theta'',z)$, since this guarantees that continuation allocations are measurable with respect to the public history. We exclude this condition here only for expositional ease, and the constraint has no bearing on our results.

z). Therefore, the welfare of households does not vary with V in this range.

4 Benchmarks

In this section, we highlight some features of the equilibrium under full information in which constraint (18) is ignored and we describe the equilibrium under full commitment in which constraint (19) is ignored. Analysis of these benchmarks allows us to highlight how our results are driven by the interaction of these two constraints. Before describing these two benchmarks, it is also useful to describe the best equilibrium, which is the equilibrium which ignores the promise keeping constraint (15) and both sustainability constraints, (18) and (19).

4.1 Best Equilibrium

Let i^* correspond to the solution to $f'(i^*) = 1$, in other words, the level of investment which equates the marginal benefit to the marginal cost of investment. Throughout the draft, we will refer to a situation in which $i_t \neq i^*$ as a distortion to production at t. It is straightforward to see that the best equilibrium for the household sets $i_t = i^*$ for all t, so that investment is efficient, and $x_t = 0$ for all t, so that policymakers receive zero rents.

We make the following assumption regarding $\underline{V}$ to guarantee that sustainability constraints (18) and (19) bind in the best sustainable equilibrium.

Assumption 1 ***(political constraints matter)*** *$\underline{V}$ satisfies*

$$\frac{v(0)}{1-\beta} < v\left(f(i^*) + \theta^N\right) + \beta\underline{V}. \tag{22}$$

Assumption 1 guarantees that the best equilibrium for the households is not politically sustainable. To see why, note that the left hand side of (22), which is the welfare of receiving zero rents forever, corresponds to the highest possible continuation value to an incumbent in power under the best equilibrium for households. The right hand side of (22) corresponds to the welfare which the incumbent could achieve under the best equilibrium for households by taxing all of output under the highest level of θ_t and being punished by replacement immediately after. Assumption 1 implies that sustainability condition (19) is not satisfied under the best equilibrium for households, which means that the best sustainable equilibrium cannot coincide with the best equilibrium. Note that Assumption 1 is trivially satisfied if $v(0) = \underline{V}(1-\beta)$, so that the policymaker's value of being thrown out of office is no worse than that associated with receiving zero rents forever.[21]

A natural question of course regards the existence of a sustainable equilibrium with positive investment, since it is clear that in a one-shot version of our model that investment by

[21] As an aside, note that even if Assumption 1 were violated, it is still potentially the case that the best equilibrium for the households is not sustainable since sustainability constraint (18) may be violated.

households is zero since they expect the incumbent to tax them one hundred percent. We make the following assumption on the discount factor which guarantees the existence of such an equilibrium for the remainder of our analysis.

Assumption 2 ***(high enough discount factor)*** *β satisfies*

$$v\left(f\left(i^{*}\right)-i^{*}+\theta^{1}\right)+\beta\frac{\sum_{n=1}^{N}\pi\left(\theta^{n}\right)v\left(f\left(i^{*}\right)-i^{*}+\theta^{n}\right)}{1-\beta}>v\left(f\left(i^{*}\right)+\theta^{1}\right)+\beta\underline{V}. \tag{23}$$

Under Assumption 2, there exists a simple stationary equilibrium in which the policymaker remains in power forever and chooses a constant tax which is independent of the shock and which leaves households indifferent between investing 0 and investing the efficient level i^{*}. The below lemma proves the existence of an equilibrium with positive investment, and we include the proof in the text since this example is useful in the discussion of equilibrium dynamics.

Lemma 1 *A sustainable equilibrium exists.*

Proof. Define δ as follows. For all (q_t, z_t), let $P_t(q_t, z_t) = 1$, $i_t(q_t, z_t) = i^{*}$, $c_t(q_t, z_t, \theta_t) = \omega$, and $x_t(q_t, z_t, \theta_t) = f(i^{*}) - i^{*} + \theta_t$ for all θ_t. The allocation satisfies (16), (17), and (20). It also implies that $V_t(q_t) = \sum_{\theta_t \in \Theta} \pi(\theta) v(f(i^{*}) - i^{*} + \theta_t) / (1-\beta) > v(0)/(1-\beta)$ for all q_t and that $x_t(q_t, z_t, \theta_t) = x_t(q_t, z_t, \widehat{\theta}) + \theta_t - \widehat{\theta}_t$ for all (q_t, z_t, θ_t) and $\widehat{\theta}_t$. Therefore, (18) is satisfied. Moreover, by Assumption 2, (19) is satisfied if $\theta_t = \theta^1$. Given the concavity of $v(\cdot)$,

$$v\left(f\left(i^{*}\right)+\theta^{n}\right)-v\left(f(i^{*})-i^{*}+\theta^{n}\right)<v(f(i^{*})+\theta^{1})-v(f\left(i^{*}\right)-i^{*}+\theta^{1})$$

for all $n > 1$, which together with Assumption 2 implies that (19) is satisfied if $\theta_t = \theta^n$. Therefore, δ is supported by sustainable equilibrium strategies. ■

An important implication of Lemma 1 is that an equilibrium without distortions to production and without replacement exists. This means that any distortions and turnover which occur in the best sustainable equilibrium must necessarily emerge as a consequence of optimality, and not because equilibria without distortions and turnover do not exists.

In the remainder of our analysis, we consider the implications of sustainability constraints (18) and (19). To facilitate the exposition of our results, we make the following assumption, which implies that the upper bound on taxes captured by the constraint in (17) does not bind along the equilibrium path.[22]

Assumption 3 ***(interior taxes)*** *$\omega - i^{*} < 0$.*

[22] The economic content of all of our results does not change if Assumption 3 is dropped. In this situation, all of our current results pertain to distortions due to sustainability constraints (18) and (19) which lead to $i(z) < i^{*}$ and not to distortions due to (17) which leads to $i(z) > i^{*}$.

4.2 Full Information Benchmark

We now consider the environment with full information, so that the citizens observe θ_t and x_t and they can condition replacement decisions on the shock to the economy as well as the policies chosen by the policymaker. This corresponds to the solution to (14) which ignores (18). In this situation, all deviations by the policymaker from prescribed policies are observable and punished by replacement.

Before proceeding, it is useful to define $\overline{V}$, the highest sustainable continuation value in the case of full information. Define $c^{\max}(\theta^n)$ and $x^{\max}(\theta^n)$ as the unique solution to

$$\max_{\{c(\theta),x(\theta)\}_{\theta\in\Theta}} \sum_{\theta\in\Theta} \pi(\theta) v(x(\theta)) \text{ s.t. } c(\theta) + x(\theta) = \omega - i^* + f(i^*) + \theta$$
$$\text{and } \sum_{\theta\in\Theta} \pi(\theta) u(c(\theta)) \geq u(\omega).$$

It is clear by feasibility that

$$\overline{V}(1-\beta) \leq \sum_{\theta\in\Theta} \pi(\theta) v(x^{\max}(\theta)).$$

The below assumption implies that the above weak inequality holds as an equality so that the repetition of the allocation associated with $c^{\max}(\theta^n)$ and $x^{\max}(\theta^n)$ satisfies sustainability constraints in the case of full information and provides the highest sustainable continuation value to the policymaker.

Assumption 4 ***(sustainability of highest value)***

$$v(x^{\max}(\theta^n)) + \beta \frac{\sum_{l=1}^{N} \pi(\theta^l) v(x^{\max}(\theta^l))}{1-\beta} \geq v(f(i^*) + \theta^n) + \beta \underline{V} \quad \forall n. \tag{24}$$

This assumption is only useful for the characterization of the full information case and has no bearing on our results.[23]

The below lemma characterizes the dynamics of distortions to production in this economy.

Lemma 2 ***(full information)*** *Under full information, the best sustainable equilibrium has the following properties:*

1. *Distortions emerge along the equilibrium path so that $i_t < i^*$ for some t, and*

2. *Distortions vanish in the long run so that*

$$\lim_{t\to\infty} \Pr\{i_{t+k} = i^* \ \forall k\} = 1.$$

[23] We make this assumption since it is a common assumption which is made in similar models in which there is full information; for example this assumption is isomorphic to Assumption 4 in Acemoglu, Golosov, and Tsyvinski (2008).

Proof. See Appendix. ■

The intuition behind this lemma are as follows. Along the equilibrium path, all actions by the policymaker are observable, and any deviation from prescribed policies is punished with replacement. Because the policymaker values holding office, the most efficient means of providing him with incentives to not deviate is to never replace him along the equilibrium path when he abides to prescribed policies. As such, there is never replacement.

In more detail, the intuition for the first part of the lemma is that distortions emerge along the equilibrium path in order to limit the resources which the policymaker can expropriate from households. This relaxes the limited commitment constraint (19) and allows society to pay lower rents to the policymaker. Formally, suppose it were the case that in the initial date, $i_0 = i^*$ and suppose for simplicity that $x_0 > 0$ for all (θ_0, z_0). In this situation, households could be made strictly better off by altering the allocation in a means which reduces the incumbent's welfare and strictly increases their welfare. Specifically, households can reduce their investment by $\epsilon > 0$ arbitrarily small, where this is achieved by making the tax system distortionary. This perturbation relaxes the right hand side of (19) by approximately $\epsilon v'\left(f\left(i^*\right) + \theta_0\right) f'\left(i^*\right)$. This allows for the reduction of rents to the policymaker under each shock θ_0 by approximately $\epsilon v'\left(f\left(i^*\right) + \theta_0\right) f'\left(i^*\right) / v'\left(x_0\right)$ so as to preserve (19). Household consumption conditional on (θ_0, z_0) changes by approximately

$$-\left(f'\left(i^*\right) - 1\right)\epsilon + \epsilon v'\left(f\left(i^*\right) + \theta_0\right) f'\left(i^*\right) / v'\left(x_0\right)$$

which exceeds 0 since $f'\left(i^*\right) = 1$. Therefore, distortions can make households strictly better off in the initial period.

The intuition for the second part of the lemma follows from the fact that *backloading* is optimal. Society optimally pays the policymaker more and more along the equilibrium path, and this is because this relaxes his limited commitment constraint (19) in the present as well as in the future. As such, even though distortions to production are efficient in the short term, in the long term they are inefficient since the policymaker is paid sufficiently that (19) is relaxed to the point that households can choose the efficient level of investment without being expropriated. Note that the absence of long run distortions under full information is not unique to our model, but common across a large class of full information principal-agent environments in which the agent suffers from limited commitment, as in Acemoglu, Golosov, and Tsyvinski (2008, 2010a,b), for example.

4.3 Full Commitment Benchmark

We now consider the environment with full commitment. Households do not observe θ_t and x_t, so that they can only condition their replacement decision based on their observation of policy. Nonetheless, the policymaker is constrained in his choice of policies, since his only possible deviations include choosing policies associated with some alternative shock $\widehat{\theta}_t = \theta_t$.

In other words, full expropriation is not feasible. As such, the full commitment benchmark corresponds to the solution to (14) which ignores (19).

Lemma 3 ***(full commitment)*** *Under full commitment, the best sustainable equilibrium features no distortions along the equilibrium path or in the long run so that* $i_t = i^* \ \forall t$*.*

Proof. See Appendix. ■

The intuition for this lemma is that in the presence of full commitment, the policymaker has limited discretion over taxes. Moreover, his continuation payoff from choosing different levels of taxes is independent of the current and future level of investment. Therefore, distortions to production cannot facilitate incentive provision, and they therefore never appear. Formally, suppose it were the case that $i_t \neq i^*$. Then it would be possible to instead perturb the solution by setting $\widehat{i_t} = i_t^*$ and $\widehat{c_t} = c_t + f(i^*) - i^* - f(i_t) + i_t$ and without altering any other portion of the contract. This perturbation would continue to be sustainable and would strictly increase household welfare. A natural question concerns the implication of asymmetric information captured by constraint (18) for the dynamics of rents and turnover. We turn to this question in the following section.

5 Analysis

We now consider the equilibrium in an environment in which the presence of limited commitment and asymmetric information interact. In light of Lemmas 2 and 3, we show in Section 5.1 that long run distortions to production emerge in this setting. In Section 5.2, we present sufficient conditions for long run turnover, and in Section 5.3, we characterize long run dynamics.

5.1 Long Run Distortions to Production

The main result of our paper is expressed in the below proposition.

Proposition 2 ***(long run distortions)*** *The best sustainable equilibrium has the following properties:*

1. *Distortions emerge along the equilibrium path so that* $i_t < i^*$ *for some* t*, and*

2. *Distortions never vanish in the long run so that*

$$\lim_{t \to \infty} \Pr\{i_{t+k} = i^* \ \forall k\} = 0.$$

Proof. See Appendix. ■

Proposition 2 states that distortions emerge and never disappear, even in the long run. This result is in stark contrast to that in Lemmas 2 and 3, and it highlights the fact that distortions emerge as a consequence of the joint interaction of the limited commitment and the asymmetric information frictions.

Intuitively, this result is due to the fact that a policymaker is always provided with dynamic incentives to not privately rent-seek, even in the long run. More specifically, if θ_t is low (high) so that shock tightens (slackens) the budget constraint and observed taxes are high (low), then the policymaker is punished (rewarded) in the future with lower (higher) payment. This ensures that the policymaker does not privately rent-seek when θ_t is high. Eventually a long sequence of negative shocks push payments to the policymaker sufficiently down that the policymaker becomes tempted to fully expropriate the investment of households. Anticipating this threat, households invest less, so that distortions to production eventually emerge as a means of preventing full expropriation.

To get a sense of the proof of the argument, note that similar arguments to those used in Section 4.2 imply that (19) binds and that there are distortions to production during an incumbent's first period of power (i.e., if $V = V_0$). Now suppose that in the long run, the commitment constraint (19) is slack, as in the case of full information, so that investment is efficient. This would require that the continuation value in the long run strictly exceed V_0 so that distortions never emerge. Suppose for simplicity that $J(V)$ is strictly concave. The first order condition with respect to $V'(\theta, z)$ together with the Envelope condition implies that the dynamics of continuation values satisfy

$$J'(V) \leq \sum_{\theta \in \Theta} \pi(\theta) J'\left(V'(\theta, z)\right), \tag{25}$$

so that the shadow marginal cost of providing a continuation value to the incumbent is a martingale. Intuitively, the continuation value in the future and future rents must weakly rise if θ_t is high as a reward for the policymaker, and they must weakly fall if θ_t is low as a punishment for the policymaker. Since $J'(V)$ is a submartingale and it is bounded from above by zero, it must converge, implying that the value of V must converge. Nonetheless, one can show that this implies suboptimal risk-sharing and is inefficient, leading to a contradiction. This therefore implies that the limited commitment constraint (19) cannot be slack in the long run.

As an example, suppose that it were the case that the equilibrium converged to a continuation value with a stationary allocation described in Lemma 1. Moreover, for simplicity, suppose there are two shocks θ^1 and θ^2 which occur with probability $1/2$. In such an equilibrium, the policymaker consumes $f(i^*) - i^* + \theta_t$ in every period and remains in power forever. Households consume ω in every period. Consider the following perturbation from this equilibrium starting from some date t. Suppose that the policymaker's consumption is increased by $\epsilon > 0$ arbitrarily small at date t if state 1 occurs at date t. Moreover, suppose that the policymaker's consumption is reduced by $.5\left(v'\left(f(i^*) - i^* + \theta^1\right)/v'\left(f(i^*) - i^* + \theta^2\right) - 1\right)\epsilon$ at date t if state 2 occurs at date t. Finally, suppose that the policymaker's consumption is reduced by $((1-\beta)/\beta)\epsilon$ at all dates and all states $t+k$ for $k \geq 1$ if state 1 occurs at date t. The policymaker's consumption at all dates $t+k$ for $k \geq 1$ if state 2 occurs at date t is unchanged. It can be verified that the proposed perturbation provides the same continuation value to the policymaker and continues to satisfy incentive compatibility. Moreover, the expected change in household welfare equals

$$\frac{u'(\omega)}{2}\left(\frac{v'\left(f\left(i^*\right)-i^*+\theta^1\right)}{v'\left(f\left(i^*\right)-i^*+\theta^2\right)}-1\right)\epsilon > 0, \tag{26}$$

which is strictly positive given the strict concavity of $v(\cdot)$. In other words, the cost to households of a decrease in consumption at date t if state 1 occurs at t is perfectly outweighed by the benefit to households of an increase in consumption at all dates $t+k$ for $k \geq 1$ if state 1 occurs at t. This means that the change in household welfare equals the increase in consumption at date t if state 2 occurs at date t.

Intuitively, the proposed stationary allocation is inefficient since the policymaker bears all the risk associated with the economic shock. A perturbation in policies which shares this risk with the households and which provides dynamic incentives to not privately rent-seek strictly increases the welfare of households. The argument relies crucially on the risk aversion on the side of the policymaker. If it were the case for example that the policymaker were risk neutral, then the term inside (26) would be equal to zero, so that there is no benefit to the perturbation and convergence to a stationary allocation without distortions would be optimal.[24,25] More generally, this argument implies that it is not possible for the continuation value to converge to *any* V, and forward iteration on this argument implies that the policymaker's continuation value must continue to decline to the minimum with positive probability after a sufficiently high number of consecutive low shocks. Nonetheless, such a reduction in continuation value eventually leads to a situation in which (19) binds, and once this happens, by the same arguments as those of Lemma 2, distortions emerge in order to prevent expropriation.

There are three important points to keep in mind in interpreting the result behind Proposition 2. First, the presence of distortions in the long run does not emerge as a consequence of the non-existence of equilibria without distortions. As Lemma 1 makes clear, such equilibria exist, but Proposition 2 states that they are inefficient.[26]

[24]We do want to note, however, that the concavity of the policymaker's welfare need not only be interpreted in terms of his preferences. Without loss of generality, one can easily interpret risk aversion on the side of the policymaker as concavity in the rent production technology in an environment with a risk neutral policymaker.

[25]A natural question concerns how our results would change if policymakers could smooth their consumption by privately borrowing and lending abroad. Allowing for this possibility in our setting would significantly complicate our analysis since it would require us to focus on a small open economy and consider the dynamics of sovereign debt. Our preliminary analysis suggests that distortions in such a setting would emerge even in the absence of asymmetric information. Details available upon request.

[26]Note that in the absence of Assumption 2 which guarantees Lemma 1, Proposition 2 continues to hold, though the reasoning for this now relies on the fact that a stationary sustainable allocation without distortions does not exist.

Second, Proposition 2 holds for any arbitrarily small variance in the private information of the policymaker. Suppose for example that $\theta_t = \{\theta^* - \sigma, \theta^* + \sigma\}$ for some $\theta^* > \sigma > 0$, where each state occurs with probability $1/2$. In this circumstance, distortions persist in the long run even for σ arbitrarily close to 0. Nevertheless, if $\sigma = 0$, then households can effectively deduce the level of rent-seeking by observation of their own consumption, so that Lemma 2 applies and distortions vanish in the long run. Therefore, the introduction of any arbitrarily small amount of privately observed uncertainty to the full information benchmark leads to the presence of long run distortions, completely altering the predictions of the full information benchmark.

Finally, the reasoning behind this proposition relies in large part on the presence of a participation constraint on the side of the households captured by (20). In the absence (20), one could construct a sustainable stationary allocation in which households consume zero and the policymaker consumes rents equal to $\omega - i^* + f(i^*) + \theta_t$ in every period. Under such an allocation, it would not be possible to perturb the equilibrium so as to induce more risk sharing between the policymaker and the households since household consumption cannot decline.

This final point elucidates the connection behind our result and that of Thomas and Worrall (1990) and Atkeson and Lucas (1992) who show that in a model of consumption risk sharing with private information, the agent's utility always declines to a minimum level. Their environment is isomorphic to our environment if constraints (17), (19), and (20) are ignored; if the households are risk-neutral; and if replacement is not allowed. As in our environment, they find that the agent's continuation value never converges to a maximal stationary level. Nonetheless, the reasoning for their result is different from ours. In our environment, this is true because even though the agent's welfare reaches the maximal level $\overline{V}$ along the equilibrium path, it must decline below $\overline{V}$ with positive probability, and this follows from optimal risk sharing. In their environment, the maximal level $\overline{V}$ is an absorbing state–much like it would be in our environment if constraint (20) were ignored–however the equilibrium never converges to such a state and this is a consequence of the Inada conditions on preferences.[27]

5.2 Long Run Turnover

In this section, we consider the dynamics of political turnover.

Proposition 3 ***(long run turnover)*** *If the set of shocks $\Theta \equiv \{\theta^1, \ldots, \theta^N\}$ is such that*

$$v(0) + \beta \frac{\sum_{n=1}^{N} \pi(\theta^n)\, v(\theta - \theta^1)}{1-\beta} > v\left(f(i^*) + \theta^1\right) + \beta \underline{V}, \tag{27}$$

[27]For example, in our environment, even if (20) were ignored so that $\overline{V}$ were associated with zero consumption for the households, the equilibrium would never converge to $\overline{V}$ if it were the case that $\lim_{V \to \overline{V}} J'(V) = -\infty$, and this would always be true if $\lim_{c \to 0} u'(c) = \infty$. Intuitively, maximally rewarding the policymaker is infinitely costly on the margin, so the equilibrium never converges to the maximal reward to the policymaker.

then the best sustainable equilibrium features long run turnover so that

$$\lim_{t\to\infty} \Pr\{P_{t+k} = 1\ \forall k\} = 0.$$

Proof. See Appendix. ■

This proposition states that if condition (27) holds, which is always true if the variance of private information is sufficiently large, then there is political turnover both along the equilibrium path and in the long run. In other words, a permanent dictator never emerges. This is because, if the variance of private information is large, then the policymaker has high private rent-seeking opportunities, and replacement is a useful means of preventing private rent-seeking.

More specifically, society has two tools for providing incentives to policymakers to not privately rent-seek. On the one hand, society can directly pay higher future rents to reward policymakers who chooses low taxes today. Though this costs societal resources, it reduces the policymaker's incentives to fully expropriate households since he values preserving power, and it allows households to choose the efficient level of investment today. On the other hand, society can instead punish policymakers who choose high taxes by removing them from office in the future. This does not cost any societal resources directly, but it raises a policymaker's incentives to fully expropriate households today since the horizon of the policymaker is reduced. In response, households are forced to invest less today, causing economic distortions. If the variance of private information is large, then a policymaker has high private rent-seeking opportunities, and providing incentives to the policymaker via payments alone is extremely costly. In this situation, the use of replacement is efficient–despite its effect on increasing economic distortions–as it allows society to make smaller payments to the policymaker.

The heuristic proof of this argument is as follows. Suppose it were the case that a permanent dictator emerged in equilibrium. Since a permanent dictator can always privately choose the policies associated with $\theta_t = \theta^1$, the informational constraints in (18) imply that the continuation welfare of such a policymaker conditional on $\theta_t = \theta^1$ must weakly exceed the left hand side of (27). Since this continuation value strictly exceeds the right hand side of (27), this implies that the limited commitment constraint (19) never binds under $\theta_t = \theta^1$. One can easily show that if this is the case, then the concavity of $v(\cdot)$ together with (18) guarantees that this constraint never binds under *any* θ_t. Then, (19) is always slack under such a permanent dictator. However, if this is the case, there are no long run distortions, so that $\lim_{t\to\infty} \Pr\{i_{t+k} = i^*\ \forall k\} > 0$, violating Proposition 2. Conceptually, whenever the constraint in (19) is slack, it implies that the continuation value to the incumbent must decline with positive probability, where this follows from (25) and the arguments in the previous section. These declines in continuation value can entail a reduction in rents. Eventually, however, because of limited liability, reductions in rents alone cannot continue to reduce the welfare of the incumbent, and turnover must be used.

Note that this result effectively generalizes the endogenous turnover result of Ferejohn (1986) to an economy in which production is determined by optimizing households and where policymakers and citizens choose fully history dependent strategies associated with the best sustainable equilibrium. Ferejohn (1986) considers an environment in which a policymaker can only be punished or rewarded with replacement and in which citizens choose Markovian strategies. The presence of turnover in his environment does not require a sufficiently large variance in the private information of the policymaker, and this is because the model does not allow for endogenous production or distortions.

More specifically, the full commitment benchmark of Section 4.3 is isomorphic to an economy with exogenous production since the limited commitment constraint (19) is ignored. In such an economy, long run turnover occurs for *any* arbitrarily small variance in the private information of the policymaker. What Proposition 3 makes clear is that long run turnover requires this variance to be sufficiently large once the limited commitment constraint (19) is taken into account. This is because if the variance of private information is too small, then replacement is too costly for society in terms of the economic distortions it entails to be used in equilibrium.

5.3 Long Run Dynamics

In this section we explore the transitional dynamics in our model. Propositions 2 shows that the model produces long run distortions and Proposition 3 shows that it produces long run turnover if the variance of shocks is sufficiently high. The below proposition shows that the model also produces long run dynamics in investment and policies. Note that since policies determine rents through (4), and these can vary with respect to the shock θ_t, we let $x_t(\theta)$ correspond to the value of rents at t conditional on the realization of the shock $\theta_t = \theta$. It is clear that if there are long run dynamics in $x_t(\theta)$, then there are also long run dynamics in policies.

Proposition 4 ***(long run dynamics)*** *If $N = 2$ or if $N > 2$ and β is sufficiently high so as to satisfy*

$$\frac{\sum_{n=1}^{N} \pi(\theta^n) v\left(f\left(i^*\right) - i^* + \theta^n\right)}{1-\beta} > v(f\left(i^*\right) + \theta^1) + \beta \underline{V} + \Gamma \tag{28}$$

for

$$\Gamma = \sum_{n=2}^{N} \left(\sum_{l=2}^{N} \pi(\theta^l) \right) \left(v(f\left(i^*\right) + \theta^1 + \theta^n - \theta^{n-1}) - v(f\left(i^*\right) + \theta^1) \right) \tag{29}$$

then best sustainable equilibrium features long run dynamics in investment and policies so that

$$\lim_{t\to\infty} \Pr\left\{i_{t+k} = i_{t+k-1}\ \forall k\right\} = 0 \text{ and } \lim_{t\to\infty} \Pr\left\{x_{t+k}\left(\theta\right) = x_{t+k-1}\left(\theta\right)\ \forall k, \theta\right\} = 0.$$

Proof. See Appendix. ■

Proposition 4 states that if $N = 2$ or if $N > 2$ but the discount factor is sufficiently large so as to satisfy (28), then there are long run dynamics in investment and rents.[28] Proposition 4 is a direct result of the fact that dynamic incentives are always provided in the long run. What Proposition 4 implies is that there is history-dependence in the sequence of investment and policies. In other words, even though shocks are i.i.d., investment and policies respond persistently to shocks. Note that these long run dynamics are significantly different relative to those in an environment with full information, since in such an environment, rents are i.i.d. and there are no distortions in the long run.

In order to further investigate the long run dynamics of our model, we perform a numerical simulation. Note that because the constraint set represented by (15) − (21) is not necessarily convex (conditional on z), a complete analytical characterization of equilibrium dynamics is not possible, and for this reason, we appeal to a numerical exercise to describe these long run dynamics. This exercise helps to provide additional intuition for the results of the previous section and also makes additional predictions. We consider the following functional forms

$$u(c) = c^{\sigma_u}; \quad v(x) = x^{\sigma_v}; \quad f(i) = A \cdot i^{\vartheta}.$$

In our simulation we choose the following parameters:

Table 1: Benchmark parameters for simulations

β	$\sigma_u = \sigma_v$	ϑ	A	ω	(θ^1, θ^2)	$\pi(\theta^1)$	$\underline{V}$
0.5	0.5	0.8	1.5	2.5	$(1.0, 1.5)$	0.5	-3.5

Figure 1 depicts the policy functions conditional on the state variable V, the continuation value promised to the policymaker. Panel A depicts the retention probability as a function of the incumbent's continuation value. It shows that an incumbent policymaker is only replaced if his promised continuation value is between $\underline{V}$, the value of being thrown out of power, and V_0 the value provided to an incumbent in his first period of power, where this probability of replacement increases as V declines in this region. The intuition for this is that it is only efficient for households to replace a policymaker if his promised value is sufficiently low since replacement serves as a punishment for the policymaker.

Panel B depicts the level of taxes as a function of the continuation value. As a reminder, note that higher taxes corresponds to higher political rents and lower household consumption. Note that the policymaker and the households share risk: both consume more during the high shock and both consume less during the low shock, and taxes are lower during the high shock and higher during the low shock. As the continuation value to the incumbent rises, taxes rise

[28] Note that condition (28) is implied by Assumption 2 if $N = 2$. The condition guarantees that the solution admits $i(z) = i^*$ if $V = \overline{V}$ so that there are no distortions whenever the continuation value approaches $\overline{V}$.

since his rents under both the high and low shock also rise.

Panel C depicts the level of investment as a function of the continuation value. It shows that distortions emerge only if the continuation value is low (i.e., the level of investment is depressed below the efficient level only if the policymaker's welfare is low). The reason behind this is that if the policymaker's welfare is low, then the value he places on remaining in power is low. Therefore it is difficult to provide him with incentives to not fully expropriate households, and for this reason, investment must be low so as to reduce the number of resources under his control and to reduce his temptation to expropriate. As his continuation value rises, it becomes possible for households to invest closer to the efficient amount while continuing to satisfy the incentive compatibility constraints on the policymaker.

Panel D shows how the policymaker is induced to choose the appropriate level of taxes and to not private rent-seek. It depicts the continuation value in the future as a function of the continuation value today. It shows that if the high shock occurs today, the policymaker is rewarded in the future with an increase in continuation value whereas if the low shock occurs today, the policymaker is punished in the future with a decrease in continuation value.

These figures provide a graphical representation for the long run dynamics of our model. If a policymaker experiences a negative economic shock, his continuation value declines, and if he experiences a positive economic shock, his continuation value increases. These dynamic incentives induce the policymaker to not privately rent-seek. Note that a decline in continuation value implies a weakly lower investment, weakly lower taxes and rents, and weakly shorter tenure. In contrast, a positive economic shock can be followed by weakly higher investment, weakly higher taxes and rents, and weakly longer tenure. These dynamics exhibit history-dependence since investment, taxes, and turnover depend on the entire history of shocks through the implied continuation value to the incumbent. Note that if a policymaker experiences a long enough sequence of low shocks, he is necessarily replaced with some probability. Importantly, periods of potential turnover are periods in which taxes are lowest (and actually negative in the simulation) and investment distortions are the highest.

We additionally consider what our model implies regarding the relationship between the turnover rate and the tenure length of policymakers. Figure 2 depicts the Kaplan-Meier estimate of the probability of replacement as a function of the tenure length of the incumbent.[29] The relationship is negative. In other words, policymakers with very short tenure are the most likely to be thrown out of power. The economic reasoning is as follows. A young incumbent is likely to have a low continuation value and be likely to be unlucky and be thrown out of power. In contrast, an older incumbent, who has lasted for several periods, is likely to have experienced many positive shocks and is thus more likely to have a high continuation value and be forgiven by citizens following a negative shock. Therefore, older incumbents experience less frequent turnover. This prediction differs from that in the model Ferejohn (1986) in

[29]The figure displays the smoothed hazard function using a rectangular smoothing kernel. Details available upon request.

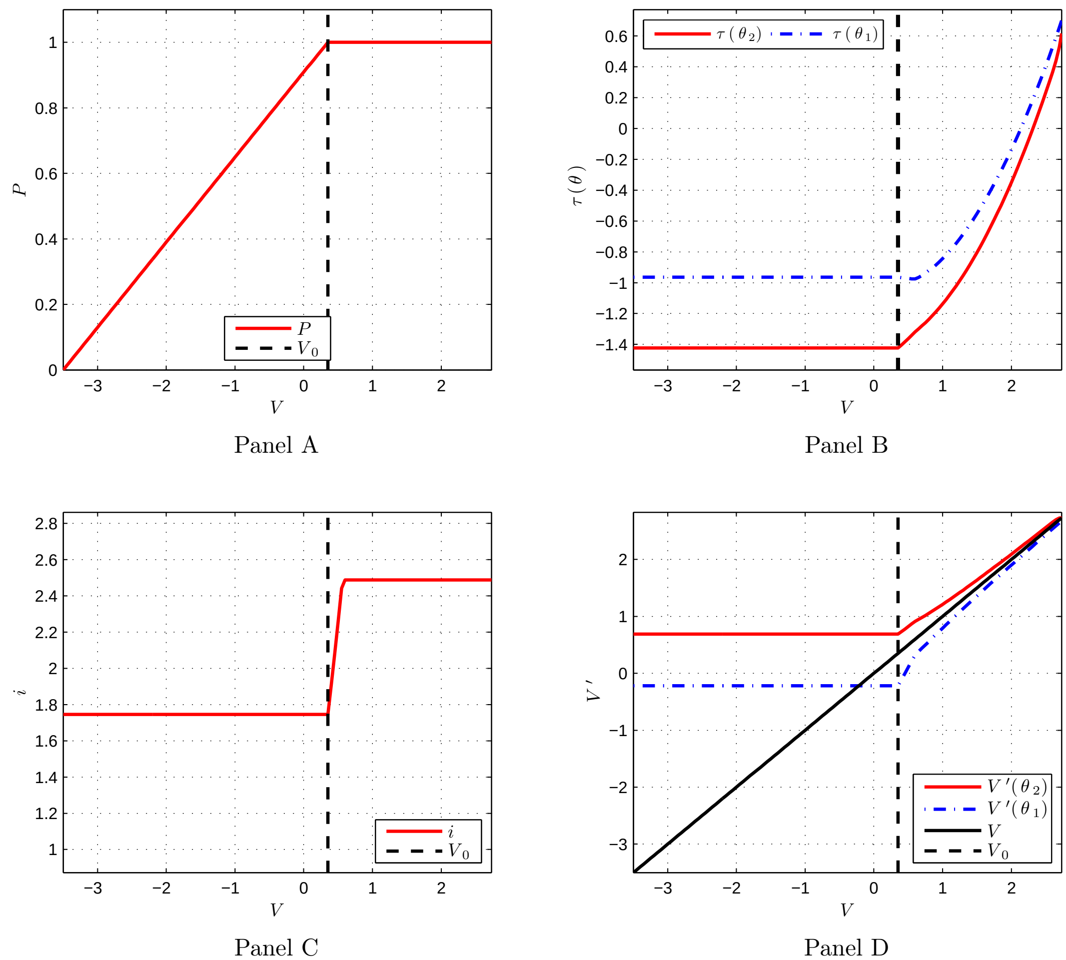

Figure 1: Policy Functions

which the replacement probability is constant and independent of tenure length.

6 Empirical Evidence

While the focus of our paper is on our theoretical results, we consider in this section the extent to which the model is consistent with the empirical patterns on the relationship between political and economic cycles. Given that our model is very stylized, we focus on discussing the qualitative implications of the model and leave a complete treatment of its quantitative implications to future work.

The model suggests that policymakers are punished for negative economic shocks with shorter tenure and with lower rents. This pattern is consistent with the previous evidence

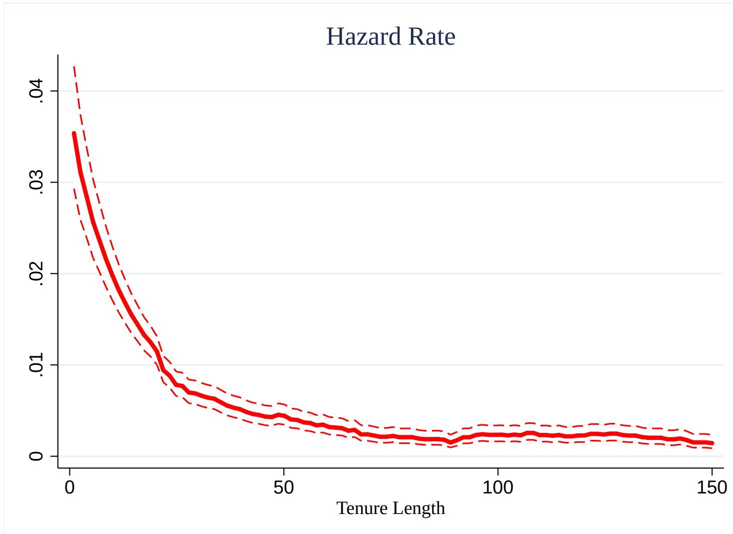

Figure 2: Smoothed hazard rate over time with 95% confidence intervals.

which suggests that policymakers are kept or replaced in response to economic shocks (e.g., Fair (1978), Lewis-Beck (1990), Achen and Bartels (2004), Wolfers (2007), and Deaton and Miller (1995)). As is the case in the model, it is often argued that these shocks are beyond the control of the policymaker, so that policymakers are effectively rewarded if they are lucky and punished if they are unlucky. In addition, Tella and Fisman (2004) find that policymakers receive a pay increase whenever taxes decrease and whenever income increases. This is also consistent with the predictions of the model.

We supplement this previous work by considering the effect of commodity price shocks in developing countries. As has been well documented, commodity price shocks are an important source of business cycles in developing countries (e.g., Deaton (1999)). In terms of our model, commodities are a key funding source for many governments in developing countries, so that global shocks to commodity prices can tighten or slacken the government budget constraint, and can do so for reasons beyond the control of policymakers.[30] An additional reason for our focus on developing countries is that political turnover in these poorly institutionalized settings is not consistently determined by regularly held elections, but can often occur through coups, revolutions, or civil wars, and this is in line with the fact that turnover can occur in any period in our model.

We begin by considering the relationship between economic growth and turnover before considering the impact of commodity price shocks. Using the data for non-OECD countries

[30] Moreover, officials may have a much better sense than citizens regarding the implications of these shocks for the tightness of the budget, and they can use the informational advantage for rent-seeking. As an example, Caselli and Michaels (2009) report that large oil output tends to be associated with an increase in instances of alleged illegal activities by mayors in Brazilian municipalities.

from Bazzi and Blattman (2011), we estimate the following baseline equation:

$$Turnover_{it} = \alpha EconomicGrowth_{it} + \eta_i + \eta_t + \epsilon_{it}. \quad (30)$$

i indexes the country and t indexes the year of the observation. $Turnover_{it}$ is a 0/1 dummy variable which takes a value of 1 if a leadership transition takes place in country i in year t (i.e., the identity of the leader in year $t+1$ is not the same as in year t). $EconomicGrowth_{it}$ is the change in log income per capita between year $t-1$ and t. η_i is a country fixed effect which controls for the country-level propensity for turnover and economic growth and η_t is a time fixed effect which controls for global trends in turnover and economic growth. ϵ_{it} is an error term.[31] The prediction of the model is that $\alpha < 0$; higher economic growth reduces the chances of turnover.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Second Stage	*Dependent Variable is Political Turnover $_t$*					
Economic Growth $_t$	-0.640	-0.540	-0.562	-0.542	-4.313	-4.074
	(0.103)	(0.104)	(0.112)	(0.116)	(2.450)	(2.537)
Economic Growth $_{t-1}$		-0.178	-0.161	-0.176		-1.617
		(0.102)	(0.105)	(0.098)		(2.039)
Economic Growth $_{t-2}$			-0.055			
			(0.098)			
Economic Growth $_t$ x Economic Growth $_{t-1}$				-0.066		
				(0.734)		
Panel B: First Stage	*Dependent Variable is Economic Growth $_t$*					
Price Shock $_{t-1}$					0.003	0.003
					(0.001)	(0.001)
Price Shock $_{t-2}$						0.000
						(0.001)
Panel C: First Stage	*Dependent Variable is Economic Growth $_{t-1}$*					
Price Shock $_{t-1}$						0.001
						(0.001)
Price Shock $_{t-2}$						0.003
						(0.001)
Observations	3849	3739	3629	3739	3849	3739
Countries	110	110	110	110	110	110
R-squared	0.12	0.12	0.12	0.12		

Fixed effects OLS regression in columns 1-4 and fixed effect instrumental variable regression in columns 5 and 6. All columns include country dummies and year dummies, and robust standard errors clustered by country are in parentheses. Base sample is an unbalanced panel, with annual data from 1961 to 2004 where the start date of the panel refers to the independent variable (i.e., t=1961, so t-1=1960). Panel A is the second stage regression where the dependent variable is political turnover. Panels B and C provide the first stage regressions in columns 5 and 6 where the dependent variable is economic growth. See text for data definitions and sources.

Table 2: Economic Shocks and Political Turnover

Column 1 in Table 2 provides the estimation of (30) and reports standard errors clustered at the country level. The coefficient on economic growth is negative and significant at the 1% level. It implies that a decrease in annual economic growth by 5% is associated with an increase in the turnover probability of the policymaker by about 3.2%. Column 2 expands equation (30) to additionally include an additional lag of economic growth and it shows that turnover

[31] The measure of turnover is from Goemans, Gleditsch, and Chiozza (2009) and excludes death of leaders by natural causes. Economic growth is from WorldBank (2009). Both of these variables are described in greater details in Bazzi and Blattman (2011). The sample is from 1961 to 2004.

responds both to growth at t as well as growth at $t-1$, with both coefficients significant at the 10% level. The coefficients in column 2 imply that a decrease in economic growth by 5% for two consecutive years leads to an increase in the turnover probability of the incumbent by about 3.6%. Column 3 adds growth at $t-2$. The coefficients on growth at t and $t-1$ are largely unchanged, and the coefficient on growth at $t-2$ is insignificant, but takes on the anticipated negative sign. Finally, column 4 considers the interaction effect between growth at t and growth at $t-1$. Though the interaction effect is insignificant, it takes on the anticipated negative sign.

In sum, the data is consistent with the model's prediction that policymakers are more likely to be thrown out of office following negative economic performance and that the effect of negative economic performance in one year can have persistent effects into future years. Columns 5 and 6 explore the extent to which commodity price shocks can drive the reduction in growth which in turn affects turnover. We do this by performing the following instrumental variable regression. Specifically, we estimate the following first stage equation:

$$EconomicGrowth_{it} = \beta PriceShock_{it-1} + \gamma_i + \gamma_t + \varepsilon_{it}. \quad (31)$$

$PriceShock_{it-1}$ is the commodity price shock for country i at date $t-1$ constructed by Bazzi and Blattman (2011). [32] γ_i and γ_t correspond to country fixed effects and time fixed effects, respectively, and ε_{it} is the error term. After estimating (31), we perform the second stage regression which replaces $EconomicGrowth_{it}$ in (30) with the predicted value of economic growth from equation (31). As such, this instrumental variable regression allows us to explore the extent to which the component of economic growth due to commodity price shocks is associated with turnover.[33,34]

Column 5 presents the results from this estimation. Panel B shows that the commodity price shock is positively associated with economic growth, and this is consistent with previous evidence such as Deaton (1999). The second stage regression in column 5 of panel A shows that the coefficient on economic growth is negative and significant at the 10% level. It implies that a decrease in annual economic growth by 5% is associated with an increase in the turnover probability of the policymaker by about 21.6%. This result is consistent with the idea that

[32] It is the log change in a country specific commodity price index which is multiplied by a time-invariant ratio of commodity export value to GDP for country i. The authors take this ratio in the years 1978-1982 which is the midpoint of the sample. The price index represents the geometric average of all commodity export prices, where the price is from international markets and each price is weighted by its lagged share in total national exports for country i. See Bazzi and Blattman (2011) for more details.

[33] We make no claim that the only means through which commodity price shocks affect turnover is through economic growth. Our objective is to document the correlation between turnover and the component of economic growth due to the commodity price shock.

[34] Our results do not change if we additionally include the contemporaneous price shock on the right hand side of (31), though it continues to be the case that the price shock at $t-1$–and not the price shock at t–which is much more tightly related to economic growth at t. This observation is consistent with results from previous related research such as in Deaton and Miller (1995).

commodity price shocks which are beyond the control of leaders have an effect on income which in turn affects the tenure of leaders. Note that the much larger coefficient in column 5 compared to the OLS estimate in column 1 suggests that leadership transitions may be more responsive to changes in growth rates due to commodity price shocks compared to changes in growth rates due to other factors. Column 6 performs the instrumental variable estimation analogue of column 2 and finds similar results, though the coefficients are much less precisely estimated. Economic growth continues to be responsive to commodity price shocks at different lags and changes in economic growth due to commodity price shocks are negatively associated with turnover.[35] Thus, turnover appears to respond to economic growth and to commodity price shocks in a persistent manner in the data as in the model.

	(1)	(2)	(3)
Panel A	*Dependent Variable is Investment Growth $_t$*		
Investment Growth$_{t-1}$	0.014	0.010	0.015
	(0.043)	(0.043)	(0.037)
Economic Growth$_t$			1.394
			(0.169)
Turnover $_t$	-0.056	-0.060	-0.037
	(0.012)	(0.012)	(0.010)
Turnover $_{t-1}$		-0.017	-0.013
		(0.012)	(0.011)
Turnover$_{t+1}$		-0.033	-0.024
		(0.009)	(0.008)
Observations	3689	3588	3366
Countries	101	101	101
R-squared	0.06	0.07	0.21
Panel B	*Dependent Variable is Tax Revenue Growth $_t$*		
Tax Revenue Growth$_{t-1}$	0.019	0.014	-0.069
	(0.063)	(0.065)	(0.066)
Economic Growth$_t$			1.128
			(0.228)
Turnover $_t$	-0.030	-0.030	-0.002
	(0.015)	(0.017)	(0.015)
Turnover $_{t-1}$		-0.028	-0.010
		(0.023)	(0.018)
Turnover$_{t+1}$		-0.016	-0.028
		(0.020)	(0.020)
Observations	485	443	441
Countries	65	60	60
R-squared	0.17	0.17	0.28

Fixed effects OLS regression in all columns. All columns include country dummies and year dummies, and robust standard errors clustered by country are in parentheses. Base sample is an unbalanced panel, with annual data from 1961 to 2004 in panel A and 1992 to 2004 in panel B, where the start date of the panel refers to the independent variable (i.e., t=1961, so t+1=1960). The dependent variable in panel A is the change in log investment and the dependent variable in panel B is the change in log tax revenue .

Table 3: Political Turnover, Investment, and Taxes.

We now explore additional predictions of the model involving the effect of turnover on economic variables. As we discussed in Section 5.3, periods of turnover are associated with the highest levels of distortions to investment and the lowest taxes. Table 3 provides suggestive

[35]Unfortunately, the instrumental variable analogue of columns 3 and 4 cannot be estimated since this involves the use of three instruments, which exacerbates the small sample bias associated with weak instruments.

evidence that this correlation holds in the data. In panel A, we regress the change in log investment on its own lag and on measures of turnover where we control for country fixed effects and time fixed effects, as in the regressions of Table 2. In column 1, we find that log investment growth is significantly lower during periods of turnover. More specifically, the coefficient, which is significant at the 1% level suggests that a period of turnover is associated with a 5.6% reduction in investment growth. Column 2 shows this pattern more generally by controlling for turnover in the following period and turnover in the previous period, and all coefficients are negative, suggesting that investment is depressed in anticipation of turnover in the following period and in response to turnover in the previous period, and this is again consistent with the predictions of the model. Finally, column 3 controls for economic growth to show that the correlation between investment growth and turnover is sustained and not driven solely by the correlation between turnover and economic growth. This evidence is consistent with the model's prediction that investment is significant decreased around periods of turnover.

In panel B we consider the effect of turnover on taxes. In this exercise we are constrained by the fact that historical data on tax revenue is limited, so that our sample size is significantly reduced by almost a factor of 10, which means that there is significantly less precision in the estimation of our coefficients.[36] In column 1, we find that periods of turnover are associated with a reduction in tax revenue growth, and this effect is significant at the 5% level. Specifically, we find that tax revenue growth is decreased by 3.0% during periods of turnover. Column 2 shows this pattern more generally by controlling for turnover in the following period and turnover in the previous period. All coefficients on turnover are negative an quantitatively large, though imprecisely estimated. Column 3 controls for economic growth and finds the results largely unchanged, though the coefficient on contemporaneous turnover is diminished. Overall, this evidence is broadly consistent with the model's prediction that taxes are significantly decreased during periods of turnover.

Our final exercise is to consider the relationship between the replacement probability and the tenure length of policymakers. Figure 2 depicts the Kaplan-Meier estimate of the probability of replacement as a function of the tenure length of the incumbent as estimated in the data.[37] It shows that the probability of replacement is decreasing in tenure length, which is consistent with the qualitative predictions of our model.

[36]An additional complication is that, in contrast to the the tax revenue variable in the model, the tax revenue measure in the data does not exclude revenue from natural resources. We therefore interpret these last set of results with this caveat in mind.

[37]For this exercise, we exclude observations of leaders who die naturally. Our results are nonetheless robust to including these observations. Details available upon request.

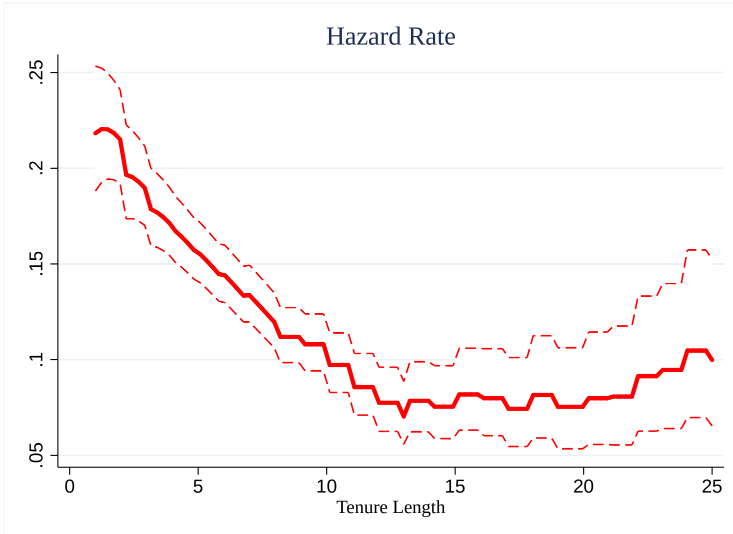

Figure 3: Smoothed hazard rate over time with 95% confidence intervals.

7 Conclusion

In this paper, we have developed a framework where political and economic cycles are jointly determined by the interaction of three frictions: the non-benevolence of policymakers, limited commitment, and asymmetric information. In our analysis, we provide conditions under which long run distortions and long run turnover emerge. In addition, our model provides predictions regarding the dynamics of tenure, investment, and taxes which are qualitatively consistent with the empirical evidence on political and economic cycles.

While we focus on a production economy with a self-interested policymaker, we believe that our results have a broader applicability to other settings. In many other interactions, a principal (represented by the citizens in our model) may be interested in providing an agent (represented by the policymaker in our model) with incentives when the agent suffers from both private information and limited commitment. As an example, consider the problem of a shareholder seeking to provide incentives to a CEO who controls the assets of the company and who privately observes its cash flows.[38] The principal must take into account two types of deviations that the agent can make for personal gain: he can privately divert cash flows and he can also sell off the company's assets for personal gain. These two frictions lead to the kind of problem which we have analyzed in this paper. In this regard, our model sheds light on dynamics of replacement and economic distortions in these other applications as well.

Our model leaves several interesting avenues for future research. First, private government information in our setting is temporary since the shocks to the government budget are i.i.d.

[38]In related work, for example, Albuquerque and Hopenhayn (2004) and Clementi and Hopenhayn (2006) consider the relationship betweeen an entrepreneur and lender-venture capitalist. However, they do not consider the joint implications of limited commitment and private information, as we do in our setup.

This assumption is not made for realism but for convenience since it maintains the common knowledge of preferences over continuation contracts and simplifies the recursive structure of the efficient sequential equilibria. Future work should consider the effect of relaxing this assumption. Second, we have assumed that all policymakers are identical, which implies that the only role for political replacement is that it incentivizes policymakers. In practice, replacement also functions as a means of selection. A natural extension of our framework would take into account both roles for replacement by allowing for multiple types of policymakers. Finally, because our economic environment is very stylized, we have not analyzed the quantitative implications of the model. A natural extension of our work is to incorporate additional layers of economic structure such as non-fully depreciating capital and sovereign debt in order to quantitatively assess the model.

References

ACEMOGLU, D., M. GOLOSOV, AND A. TSYVINSKI (2008): "Political Economy of Mechanisms," *Econometrica*, 76(3), 619–641.

——— (2010a): "Dynamic Mirrlees Taxation Under Political Economy Constraints," *Review of Economic Studies*, 77(3), 841–881.

——— (2010b): "Political economy of Ramsey taxation," *Journal of Public Economics*.

ACHEN, C., AND L. BARTELS (2004): "Blind Retrospection. Electoral Responses to Drought, Flu, and Shark Attacks," *Working Papers (Centro de Estudios Avanzados en Ciencias Sociales)*.

AGUIAR, M., AND M. AMADOR (2011): "Growth in the Shadow of Expropriation," *The Quarterly Journal of Economics*, 126(2), 651–697.

ALBUQUERQUE, R., AND H. HOPENHAYN (2004): "Optimal Lending Contracts and Firm Dynamics," *Review of Economic Studies*, 71(2), 285–315.

ALESINA, A. (1988): "Macroeconomics and Politics," *NBER Macroeconomics Annual*, pp. 13–62.

ALESINA, A., AND R. PEROTTI (1994): "The Political Economy of Budget Deficits," Discussion paper, National Bureau of Economic Research.

ALESINA, A., N. ROUBINI, AND G. COHEN (1997): *Political cycles and the macroeconomy*. MIT Press.

ALESINA, A., AND G. TABELLINI (1990): "A Positive Theorv of Fiscal Deficits and Government Debt," *The Review of Economic Studies*, 57(3), 403–414.

AMADOR, M., I. WERNING, AND G. ANGELETOS (2006): "Commitment vs. Flexibility," *Econometrica*, 74(2), 365–396.

ATHEY, S., A. ATKESON, AND P. KEHOE (2005): "The Optimal Degree of Discretion in Monetary Policy," *Econometrica*, 73(5), 1431–1475.

ATKESON, A., AND R. LUCAS (1992): "On Efficient Distribution With Private Information," *The Review of Economic Studies*, 59(3), 427.

AZZIMONTI, M. (2010): "Barriers to Investment in Polarized Societies," *American Economic Review, forthcoming*.

BANKS, J., AND R. SUNDARAM (1998): "Optimal Retention in Agency Problems* 1," *Journal of Economic Theory*, 82(2), 293–323.

BARRO, R. J. (1973): "The Control of policymakers: An Economic Model," *Public Choice*, pp. 19–42.

Bates, R. (2008): *When Things Fell Apart: State Failure in Late-Century Africa*. Cambridge Univ Pr.

Battaglini, M., and S. Coate (2008): "A Dynamic Theory of Public Spending, Taxation, and Debt," *American Economic Review*, 98(1), 201–236.

Baylies, C., and M. Szeftel (1992): "The fall and rise of multi-party politics in Zambia," *Review of African Political Economy*, 19(54), 75–91.

Bazzi, S., and C. Blattman (2011): "Economic Shocks and Conflict: The (Absence of?) Evidence From Commodity Prices," *Working Paper*.

Benveniste, L., and J. Scheinkman (1979): "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica: Journal of the Econometric Society*, pp. 727–732.

Bergoeing, R., P. Kehoe, T. Kehoe, and R. Soto (2002): "A Decade Lost and Found: Mexico and Chile in the 1980s," *Review of Economic Dynamics*, 5(1), 166–205.

Besley, T. (2006): *Principled Agents?: The Political Economy of Good Government*. Oxford University Press, USA.

Caballero, R., and P. Yared (2010): "Future Rent-Seeking and Current Public Savings," *Journal of International Economics*, 82, 1124–136.

Caselli, F., and G. Michaels (2009): "Do Oil Windfalls Improve Living Standards? Evidence From Brazil," *NBER Working Papers*.

Chari, V., and P. Kehoe (1993a): "Sustainable Plans and Debt," *Journal of Economic Theory*, 61(2), 230–261.

Chari, V., and P. Kehoe (1993b): "Sustainable Plans and Mutual Default," *The Review of Economic Studies*, 60(1), 175.

Clementi, G., and H. Hopenhayn (2006): "A Theory of Financing Constraints and Firm Dynamics," *Quaterly journal of economics*, 1(1), 229–265.

Deaton, A. (1999): "Commodity Prices and Growth in Africa," *The Journal of Economic Perspectives*, 13(3), 23–40.

Deaton, A., and R. Miller (1995): "International Commodity Prices, Macroeconomic Performance, and Politics in Sub-Saharan Africa," *Princeton Studies in International Economics*.

Drazen, A. (2000): "The Political Business Cycle After 25 Years," *NBER Macroeconomics Annual*, pp. 75–117.

Egorov, G. (2009): "Political Accountability Under Special Interest Politics," *Working Paper*.

Fair, R. (1978): "The Effect of Economic Events on Votes for President," *The Review of Economics and Statistics*, 60(2), 159–173.

Fearon, J. (2010): "Coordinating on Democracy," *Quarterly Journal of Economics, forthcoming*.

Ferejohn, J. (1986): "Incumbent Performance and Electoral Control," *Public choice*, 50(1), 5–25.

Goemans, H., K. Gleditsch, and G. Chiozza (2009): "Introducing Archigos: A Dataset of Political Leaders," *Journal of Peace Research*, 46(2), 269–283.

Krusell, P., and J. Rios-Rull (1999): "On the Size of US Government: Political Economy in the Neoclassical Growth Model," *American Economic Review*, pp. 1156–1181.

Lewis-Beck, M. (1990): *Economics and elections: The Major Western Democracies*. University of Michigan Press.

Li, J., and N. Matouschek (2011): "Managing Conflicts in Relational Contracts," *Working Paper*.

Persson, T., and L. Svensson (1989): "Why a Stubborn Conservative Would Run a Deficit: Policy With Time-Inconsistent Preferences," *The Quarterly Journal of Economics*, 104(2), 325–345.

Persson, T., and G. Tabellini (2000): *Political Economics*. MIT press.

Ray, D. (2002): "The Time Structure of Self-Enforcing Agreements," *Econometrica*, 70(2), 547–582.

Rogoff, K. (1990): "Equilibrium Political Budget Cycles," *The American Economic Review*, pp. 21–36.

Simutanyi, N. (1996): "The politics of structural adjustment in Zambia," *Third World Quarterly*, 17(4), 825–839.

Sleet, C. (2001): "On Credible Monetary Policy and Private Government Information," *Journal of Economic Theory*, 99(1-2), 338–376.

——— (2004): "Optimal Taxation With Private Government Information," *Review of Economic Studies*, 71(4), 1217–1239.

Song, Z., K. Storesletten, and F. Zilibotti (2009): "Rotten Parents and Disciplined Children: A Politico-economic Theory of Public Expenditure and Debt," *Working Paper*.

Tella, R., and R. Fisman (2004): "Are Politicians Really Paid Like Bureaucrats?," *Journal of Law and Economics*, 47.

Thomas, J., and T. Worrall (1990): “Income Fluctuation and Asymmetric Information: An example of a Repeated Principal-agent Problem* 1,” *Journal of Economic Theory*, 51(2), 367–390.

——— (1994): “Foreign Direct Investment and The Risk of Expropriation,” *The Review of Economic Studies*, 61(1), 81–108.

Wolfers, J. (2007): “Are Voters Rational? Evidence from Gubernatorial Elections,” *Working Paper*.

WorldBank (2009): *World Development Indicators 2009*. The World Bank.

Appendix

A Proofs of Section 3

A.1 Proof of Proposition 1

Step 1. We begin by first proving the necessity of these conditions. (7), (8), and (11) must be satisfied by feasibility and by the fact that, in choosing their level of investment, households can always choose $i(z_t) = 0$ forever which provides them with a utility of at least $u(\omega) / (1-\beta)$. The necessity of (9) follows from the fact that conditional on (q_t, z_t, θ_t), the policymaker can choose the taxes appropriate for $(q_t, z_t, \widehat{\theta}_t)$ for $\widehat{\theta} \neq \theta_t$ and he can follow the equilibrium strategy from $t+1$ onward. From (4), this provides him with immediate rents equal to $x_t(q_t, z_t, \widehat{\theta}) + \theta_t - \widehat{\theta}$ and his continuation value from $t+1$ onward equals $V_{t+1}(q_t, z_t, \widehat{\theta})$. Condition (9) guarantees that this privately observed deviation is weakly dominated. The necessity of (10) follows from the fact that conditional on (q_t, z_t, θ_t), the policymaker can choose to tax the maximum which from (4) provides him with rents equal to $f(i_t(q_t, z_t)) + \theta_t$. Given that $v(0) \geq \underline{V}(1-\beta)$, his continuation value from $t+1$ onward following the deviation must weakly exceed $\underline{V}$. Condition (10) guarantees that this deviation is weakly dominated.

Step 2. For sufficiency, consider an allocation which satisfies (7) $-$ (11). Since feasibility is satisfied, we only need to check that there exist policies so as to induce households to choose the level of investment $i_t(q_t, z_t)$ at every (q_t, z_t). Suppose that conditional on θ_t, the government sets taxes equal to one hundred percent if a household has not chosen the prescribed investment sequence up to and including date t. Otherwise, if a household has chosen the prescribed investment level, the government sets taxes equal to $x_t(q_t, z_t, \theta_t) - \theta_t$ for each θ_t, where this is feasible given (8). Given this tax structure, any investment level for households other than $i_t(q_t, z_t)$ is strictly dominated by investing 0 forever which yields $u(\omega) / (1-\beta)$. From (11), investing $i_t(q_t, z_t)$ weakly dominates investing 0, so that the allocation satisfies household optimality.

We now verify that the allocation is sustained by equilibrium strategies by the incumbent policymaker and the representative citizen. Suppose that following a public deviation by the policymaker at t, the representative citizen replaces the incumbent at $t+1$ for all realizations of z_{t+1}. Moreover, following any public deviation by the representative citizen, the equilibrium allocations from $t+1$ onward are unchanged. We now verify that the allocation is sustainable. We only consider single period deviations since $\beta < 1$ and since continuation values are bounded. Let us consider the incentives of the policymaker to deviate. Conditional on (q_t, z_t, θ_t), the policymaker can deviate privately or publicly. Any private deviation requires the policymaker to choose policies prescribed for $(q_t, z_t, \widehat{\theta}_t)$ for $\widehat{\theta}_t \neq \theta$. This provides him with immediate rents equal to $x_t(q_t, z_t, \widehat{\theta}) + \theta_t - \widehat{\theta}$ and his continuation value from $t+1$ onward equals $V_{t+1}(q_t, z_t, \widehat{\theta})$. Condition (9) implies that this privately observed deviation is weakly dominated. Alternatively the policymaker can deviate publicly. Since all public deviations yield a continuation value $\underline{V}$ at $t+1$, the best public deviation maximizes immediate rents, and this is achieved with a one hundred percent tax. This yields rents equal to $f(i_t(q_t, z_t)) + \theta_t$ at t and a continuation value $\underline{V}$ from $t+1$ onward. Condition (10) guarantees that this deviation is weakly dominated. Now let us consider the incentives of the representative citizen to not deviate. If he deviates from the replacement decision, the continuation equilibrium is identical as if he had not deviated. As such, his welfare is independent of the replacement decision, and for this reason any deviation is weakly dominated.■

B Technical Results

In this section, we prove technical results regarding $J(V)$ which simplify our analysis.

Lemma 4 *$J(V)$ satisfies the following properties: (i) It is weakly concave in V, (ii) it satisfies $J(V) = \overline{J}$ for $V \in [\underline{V}, V_0]$ and it is strictly decreasing in V if $V \in (V_0, \overline{V}]$.*

Proof. Proof of part (i). Consider two continuation values $\{V', V''\}$ associated with corresponding solutions α' and α'' which provide welfare $J(V')$ and $J(V'')$. Define $V^\kappa = \kappa V' + (1-\kappa) V''$ for some $\kappa \in (0,1)$. It must be that

$$J(V^\kappa) \geq \kappa J(V') + (1-\kappa) J(V'') \;\; \forall \kappa \in (0,1) \tag{32}$$

establishing the weak concavity of $J(V)$. Suppose this were not the case. Define α^κ as follows:

$$\alpha^\kappa|_z = \begin{cases} \alpha'|_{\frac{z}{\kappa}} & \text{if } z \in [0, \kappa) \\ \alpha''|_{\frac{z-\kappa}{1-\kappa}} & \text{if } z \in [\kappa, 1] \end{cases},$$

where $\alpha^\kappa|_z$ corresponds to the component of α^κ conditional on the realization of z, and $\alpha'|_z$ and $\alpha''|_z$ are defined analogously. Since α' and α'' satisfy (16) − (21), α^κ satisfies (16) − (21) and it provides continuation value V^κ, achieving a welfare equal to the right hand side of (32). Therefore, (32) must be satisfied since $J(V^\kappa)$ must weakly exceed the welfare achieved under this feasible solution.

Proof of part (ii) We first prove that $J(V)$ is weakly decreasing in V. Suppose by contradiction that $J(V') < J(V'')$ for some $V'' > V'$ where V' and V'' are associated with corresponding solutions α' and α'', respectively. Define $\widehat{\alpha}'$ as follows:[39]

$$\widehat{\alpha}'|_z = \begin{cases} P(z) = 0 & \text{if } z \in [0, (V''-V')/(V''-\underline{V})) \\ \alpha''|_{\frac{z-(V''-V')/(V''-\underline{V})}{1-(V''-V')/(V''-\underline{V})}} & \text{if } z \in [(V''-V')/(V''-\underline{V}), 1] \end{cases},$$

[39]Note that if $P(z) = 0$, then the values of $i(z)$, $c(\theta, z)$, $x(\theta, z)$, and $V'(\theta, z)$ are payoff irrelevant since households receive $\overline{J}$ and the replacement policymaker receives V_0.

$\widehat{\alpha}'$ satisfies (16) − (21) and provides continuation value V' so that it satisfies (15), and it achieves household welfare equal to

$$\frac{V''-V'}{V''-\underline{V}}\overline{J}+\frac{V'-\underline{V}}{V''-\underline{V}}J\left(V''\right)\geq J\left(V''\right)>J\left(V'\right)$$

where we have used the fact that $\overline{J}\geq J\left(V\right)\forall V$ by definition. This contradicts the fact that α' is a solution to (14) − (21). Now note that $J\left(V\right)=\overline{J}$ for all $V\in\left[\underline{V},V_0\right]$ since by definition, $J\left(\underline{V}\right)=J\left(V_0\right)=\overline{J}\geq J\left(V\right)$ and since $J\left(\cdot\right)$ is weakly concave.

Note that by definition of V_0, $J\left(V\right)<J\left(V_0\right)=\overline{J}$ if $V>V_0$ since V_0 must represent the highest continuation value that the policymaker can receive conditional on the households receiving their highest continuation welfare $\overline{J}$. Given that $J\left(V\right)$ is weakly concave, this means that $J\left(V\right)$ is strictly decreasing in V if $V\in\left(V_0,\overline{V}\right]$. ■

We now move to prove the continuous differentiability of $J\left(V\right)$ for $V\in\left(\underline{V},\overline{V}\right)$. From part (ii), we only need to consider continuation values $V\geq V_0$, since otherwise $J\left(V\right)=\overline{J}$ and $J'\left(V\right)=0$. The below preliminary result implies that there is no turnover if $V\geq V_0$.

Lemma 5 *If $V\geq V_0$, then the solution to (14) − (21) admits $P\left(z\right)=1\ \forall z$.*

Proof. We first establish that $V_0>\underline{V}$. This is because from (19),

$$V_0\geq\sum_{n=1}^{N}\pi\left(\theta^n\right)v\left(f\left(i\left(z\right)\right)+\theta^n\right)+\beta\underline{V}>v\left(0\right)+\beta\underline{V}\geq\underline{V}.$$

Consider the solution α given $V\geq V_0$ and suppose that by contradiction $P\left(z\right)=0$ for some positive measure z. Define $q=\int_0^1 P\left(z\right)dz\in\left(0,1\right)$ and

$$V_q=\frac{\int_0^1 P\left(z\right)\left[\left(\sum_{n=1}^{N}\pi\left(\theta^n\right)\left(v\left(x\left(\theta^n,z\right)\right)+\beta V'\left(\theta^n,z\right)\right)\right)\right]dz}{q}.$$

V_q corresponds to the continuation value to the policymaker conditional on preserving power. It is clear that since $V\geq V_0>\underline{V}$, (15) and $q<1$ imply that $V_q>V\geq V_0$. The weak concavity of $J\left(\cdot\right)$ implies that

$$J\left(V\right)\geq\left(1-q\right)\overline{J}+qJ\left(V_q\right). \tag{33}$$

Moreover, if it were the case that (33) were a strict inequality, then this would imply that the solution α conditional on $P\left(z\right)=1$ provides a continuation value to the policymaker V_q and yields welfare to households strictly greater than $J\left(V_q\right)$, which contradicts the fact that the solution to (14) subject to $V=V_q$ is optimal. It follows that (33) holds with equality so that

$$J\left(V\right)=\left(1-q\right)\overline{J}+qJ\left(V_q\right)<\overline{J}. \tag{34}$$

Define $\widetilde{q}$ as the value which satisfies $V = (1-\widetilde{q})\,V_0 + \widetilde{q}V_q$. It is clear that $\widetilde{q} < q$ since $V_0 > \underline{V}$. The weak concavity of $J(V)$ implies that

$$J(V) \geq (1-\widetilde{q})\,J(V_0) + \widetilde{q}J\left(V_q\right) = (1-\widetilde{q})\,\overline{J} + \widetilde{q}J\left(V_q\right),$$

which contradicts (34) since $\widetilde{q} < q$. This establishes that if $V \geq V_0$, then $P(z) = 1\ \forall z$. ■

We establish the following preliminary lemmas. Define $C_{n,n+k}$ as follows:

$$C_{n,n+k} = v\left(x\left(\theta^n, z\right)\right) + \beta V'\left(\theta^n, z\right) - v(x(\theta^{n+k}, z) + \theta^n - \theta^{n+k}) - \beta V'(\theta^{n+k}, z).$$

The following lemma characterizes the set of allocations α defined in (13) which satisfy (15) − (21). This lemma simplifies the problem by illustrating which set of constraints are redundant and can be ignored in different circumstances. To simplify notation, we let $\alpha|_z$ correspond to the component of α conditional on the realization of z.

Lemma 6 *For a given allocation $\alpha|_z$, the following must be true:*

1. *If $\alpha|_z$ satisfies (18), then $x(\theta, z) - \theta$ is weakly decreasing in θ and $V'(\theta, z)$ is weakly increasing in θ,*
2. *If $\alpha|_z$ satisfies $C_{n+1,n} \geq 0$ and $C_{n,n+1} \geq 0\ \forall n < N$, or if $\alpha|_z$ satisfies $C_{n+1,n} = 0$ and $x(\theta^n, z) - \theta^n \geq x\left(\theta^{n+1}, z\right) - \theta^{n+1}\ \forall n < N$, then $\alpha|_z$ satisfies (18) $\forall\theta$,*
3. *If $\alpha|_z$ satisfies (18) $\forall\theta$ and (17) for $\theta = \theta^1$, then $\alpha|_z$ satisfies (17) $\forall\theta$,*
4. *If $\alpha|_z$ satisfies (18) $\forall\theta$, (17) $\forall\theta$, and (19) for $\theta = \theta^1$, then $\alpha|_z$ satisfies (19) $\forall\theta$,*
5. *If $\alpha|_z$ satisfies $V'\left(\theta^1, z\right) \geq \underline{V}$ and (17) holding with equality for $\theta = \theta^1$, then $\alpha|_z$ satisfies (19) for θ^1, and if $\alpha|_z$ satisfies $V'\left(\theta^1, z\right) \geq \underline{V}$ and (19) holding with equality for $\theta = \theta^1$, then $\alpha|_z$ satisfies (17) for θ^1.*

Proof. Proof of part (i). Note that the constraints that $C_{n,n+k} \geq 0$ and $C_{n+k,n} \geq 0$ for $k \geq 1$ together imply:

$$v(x(\theta^{n+k}, z)) - v(x(\theta^{n+k}, z) - (\theta^{n+k} - \theta^n)) \geq v(x\left(\theta^n, z\right) + \theta^{n+k} - \theta^n) - v\left(x\left(\theta^n, z\right)\right),$$

which given the concavity of $v(\cdot)$ can only be true if $x\left(\theta^{n+k}, z\right) - \theta^{n+k} \leq x\left(\theta^n, z\right) - \theta^n$. This establishes that $x(\theta, z) - \theta$ is weakly decreasing in θ. Given this fact, it follows that for $C_{n+k,n} \geq 0$ to hold, it is necessary that $V'\left(\theta^{n+k}, z\right) \geq V'\left(\theta^n, z\right)$.

Proof of part (ii). This is proved by induction. Suppose that $C_{n+1,n} \geq 0$ and $C_{n,n+1} \geq 0$ $\forall n < N$. From part (i), this implies that $x\left(\theta^n, z\right) - \theta^n \geq x\left(\theta^{n+1}, z\right) - \theta^{n+1}$, which given the concavity of $v(\cdot)$ implies that

$$v(x(\theta^{n+1},z)+\theta^{n+2}-\theta^{n+1})-v(x(\theta^{n+1},z))\geq$$
$$v\left(x\left(\theta^{n},z\right)+\theta^{n+2}-\theta^{n}\right)-v(x\left(\theta^{n},z\right)+\theta^{n+1}-\theta^{n}).$$

Together with the fact that $C_{n+1,n}\geq 0$ and $C_{n+2,n+1}\geq 0$, the above condition implies that $C_{n+2,n}\geq 0$. Forward iteration on this argument implies that $C_{n+k,n}\geq 0$ for all n and k for which $n+k\leq N$. Analogous arguments can be used to show that if $C_{n,n+1}\geq 0$ for all $n<N$, then $C_{n,n+k}\geq 0$ for all n and k for which $n+k\leq N$.

Now suppose that $C_{n+1,n}=0$ and $x(\theta^{n+1},z)-\theta^{n+1}\leq x(\theta^{n},z)-\theta^{n}$ $\forall n<N$. Then this implies that $C_{n,n+1}\geq 0$ for all $n<N$, and given that this is the case, the same arguments as above can be applied. To see why, suppose instead that $C_{n,n+1}<0$. Together with the fact that $C_{n+1,n}=0$, this would imply that

$$v\left(x\left(\theta^{n},z\right)\right)-v(x(\theta^{n+1},z)-(\theta^{n+1}-\theta^{n}))<v(x\left(\theta^{n},z\right)+(\theta^{n+1}-\theta^{n}))-v(x(\theta^{n+1},z)),$$

from concavity of $v(\cdot)$ the above implies that $x(\theta^{n+1},z)-\theta^{n+1}>x(\theta^{n},z)-\theta^{n}$ which is a contradiction.

Proof of part (iii). Suppose that (17) holds for $\theta=\theta^{1}$. Then given that (18) also holds, from part (i), (17) holds $\forall\theta$.

Proof of part (iv). Condition (18) for $\theta=\theta^{n}$ implies that

$$v\left(x\left(\theta^{n},z\right)\right)+\beta V'\left(\theta^{n},z\right)\geq v(x(\theta^{1},z)+\theta^{n}-\theta^{1})+\beta V'(\theta^{1},z)$$

which when combined with (19) for $\theta=\theta^{1}$ implies that

$$v\left(x\left(\theta^{n},z\right)\right)+\beta V'\left(\theta^{n},z\right)\geq v(x(\theta^{1},z)+\theta^{n}-\theta^{1})-v(x(\theta^{1},z))+v(f\left(i\left(z\right)\right)+\theta^{1})+\beta\underline{V}. \quad (35)$$

The left hand side of (35) equals the left hand side of (19) for $\theta=\theta^{n}$. The concavity of $v(\cdot)$ implies that the right hand side of (35) weakly exceeds $v\left(f\left(i\left(z\right)\right)+\theta^{n}\right)+\beta\underline{V}$ since (17) implies that $x\left(\theta^{1},z\right)\leq f\left(i\right)+\theta^{1}$.

Proof of part (v). Suppose that (17) is an equality for $\theta=\theta^{1}$. The fact that $V'\left(\theta^{1},z\right)\geq\underline{V}$ and together with (17) which is an equality implies that (19) is satisfied for $\theta=\theta^{1}$. Suppose that (19) binds for $\theta=\theta^{1}$. Since $V'\left(\theta^{1},z\right)\geq\underline{V}$, it follows that (17) is implied for $\theta=\theta^{1}$. ■

The main takeaways from Lemma 6 are as follows. The solution to the problem in (14) – (21) is the same as the solution to the relaxed problem which ignores (17) and (19) for $n>1$ and which ignores the non-local constraints in (18). In addition, if one of either constraints (17) and (19) holds with equality, then the other is made redundant.

We now establish the existence of a solution to (14) – (21) with specific properties. Our first step is to show that there is always a solution in which the downward constraints in (18) bind.

Lemma 7 *If $V \in \left[V_0, \overline{V}\right)$, there exists a solution to (14) − (21) with the property that* $C_{n+1,n} = 0$ $\forall n < N$ *and* $\forall z$.

Proof. Consider a solution to the program α for which conditional on z, $C_{n+1} > C_n$ for some n. We can show that there exists a perturbation of this solution which satisfies all of the constraints and yields weakly greater welfare to the households and for which $C_{n+1,n} = 0$ for all n. Consider an alternative solution to the program $\widehat{\alpha}$ which is identical to α with the exception that $\widehat{V}'(\theta, z)$ satisfies the following system of equations $\forall n < N$:

$$\sum_{n=1}^{N} \pi(\theta^n) \widehat{V}'(\theta^n, z) = \sum_{n=1}^{N} \pi(\theta^n) V'(\theta^n, z) \tag{36}$$

$$\widehat{V}'(\theta^{n+1}, z) = \widehat{V}'(\theta^n, z) + \left[v(x(\theta^n, z) + \theta^{n+1} - \theta^n) - v(x(\theta^{n+1}, z))\right] / \beta \tag{37}$$

We now verify that the perturbed solution satisfies all of the constraints of the program. It satisfies (16) and (17) since these are satisfied under the original allocation, and it satisfies (15) given (36) and the fact that (15) is also satisfied in the original allocation. From (37), it satisfies $C_{n+1,n} = 0$ $\forall n < N$. Moreover, it satisfies $C_{n,n+1} \geq 0$ $\forall n < N$ since if this were not the case, then together with the fact that $C_{n+1,n} = 0$, it would imply that

$$v(x(\theta^{n+1})) - v(x(\theta^{n+1}) - (\theta^{n+1} - \theta^n)) < v(x(\theta^n) + \theta^{n+1} - \theta^n) - v(x(\theta^n)),$$

which given the concavity of $v(\cdot)$ violates the fact that $x(\theta^n) \geq x\left(\theta^{n+1}\right) - \left(\theta^{n+1} - \theta^n\right)$ established in part (i) of Lemma 6. From part (ii) of Lemma 6, this implies that (18) is satisfied for all θ and $\widehat{\theta}$. From part (iv) of Lemma 6, we need only verify (19) for $\theta = \theta^1$, since (19) for other θ's are implied by the satisfaction of (17) and (18). This is implied by the fact that (36) and (37) imply that $\widehat{V}'\left(\theta^1, z\right) \geq V'\left(\theta^1, z\right)$. To see why this is true, note that the fact that $C_{n+1,n} \geq 0$ in the original solution implies that

$$\sum_{n=1}^{N} \pi(\theta^n) V'(\theta^n, z) \geq$$
$$V'(\theta^1, z) + \sum_{n=2}^{N} \pi(\theta^n) \sum_{l=1}^{n-1} (v(x(\theta^{n-l}, z) + \theta^{n-l+1} - \theta^{n-l}) - v(x(\theta^{n-l+1}, z)))/\beta$$

which combined with (36) and (37) implies that $\widehat{V}'\left(\theta^1, z\right) \geq V'\left(\theta^1, z\right)$. Analogous arguments imply that $\widehat{V}'\left(\theta^N, z\right) \leq V'\left(\theta^N, z\right)$, which together with part (i) of Lemma 6 implies that (21) is satisfied. Given (36) and (37) and the weak concavity of $J(\cdot)$, it follows that for all z,

$$\sum_{n=1}^{N} \pi(\theta^n) J(\widehat{V}'(\theta^n, z)) \geq \sum_{n=1}^{N} \pi(\theta^n) J(V'(\theta^n, z)), \tag{38}$$

since V' is a mean preserving spread over $\widehat{V}'$. Therefore, (20) is satisfied. Therefore, $\widehat{\alpha}$ satisfies all of the constraints of the problem, and by (38), it weakly increases the welfare of the households. ■

What Lemma 7 shows in light of part (ii) of Lemma 6 is that there exists a solution to (14) for which constraint (18) is replaced with the constraint that $C_{n+1,n} = 0$ and $x(\theta^n, z) - \theta^n \geq x\left(\theta^{n+1}, z\right) - \theta^{n+1}$ $\forall n < N$. Next we describe necessary properties of the solution to (14).

Lemma 8 *The solution to* (14) $-$ (21) *has the following necessary properties:*

1. *If $V = \overline{V}$, then* (20) *binds* $\forall z$,

2. *If $V \in \left[V_0, \overline{V}\right)$, then (a) $i(z) > 0$ $\forall z$, (b) $c(\theta, z) > 0$ $\forall \theta, z$, and (c)* (20) *doesn't bind for some* z.

Proof. Proof of part (i). Suppose that $V = \overline{V}$ but that (20) does not bind for some z. It is clear that conditional on z, the allocation α must provide a welfare of $\overline{V}$ to the policymaker since, otherwise it would be possible to make the policymaker strictly better by providing him the highest welfare for all z's and continuing to satisfy all of the constraints of the problem. Therefore, we can without loss of generality focus on the solution given $V = \overline{V}$ for which α is the same across z's. Moreover, by Lemma 7, we can consider such a solution for which $C_{n+1,n} = 0$ $\forall n < N$. It is clear by the arguments of Lemma 7 that if (20) is slack under some original allocation for which $C_{n+1,n} > 0$, then it continues to be slack under a perturbed allocation for which $C_{n+1,n} = 0$. There are two cases to consider.

Case 1. Suppose it were the case that $C_{n-1,n} = 0$ $\forall n \leq N$ so that $V'(\theta^n, z) = V'\left(\theta^{n-1}, z\right)$ and $c(\theta^n, z) = c\left(\theta^{n-1}, z\right)$ $\forall n < N$. Then this implies that $V'(\theta^n, z) = \overline{V}$ and $c(\theta^n, z) = 0$ $\forall n$. To see why, note that if $V'(\theta^n, z) < \overline{V}$, then it would be possible to increase $V'(\theta^n, z)$ by $\epsilon > 0$ arbitrarily small $\forall n$ while continuing to satisfy all of the constraints of the problem and making the policymaker strictly better off. Suppose instead that $c(\theta^n, z) > 0$. Then it would be possible to increase $i(z)$ by $\epsilon > 0$, increase $x(\theta^n, z)$ by $f(i + \epsilon) - f(i)$ $\forall n$, and reduce $c(\theta^n, z)$ by ϵ $\forall n$ while continuing to satisfy the constraints of the problem and making the policymaker strictly better off. However, if it is the case that $V'(\theta^n, z) = \overline{V}$ and $c(\theta^n, z) = 0$ $\forall n$, then this implies that households are receiving a consumption of 0 forever, which violates (20).

Case 2. Suppose it is the case that $C_{n-1,n} > 0$ for some $n < N$. We rule out this case by induction. Suppose that $C_{N-1,N} > 0$. Then this implies that $V'\left(\theta^N, z\right) = \overline{V}$ and $c\left(\theta^N, z\right) = 0$. This is because of analogous arguments as those of case 1. If $V'\left(\theta^N, z\right) < \overline{V}$, then $V'\left(\theta^N, z\right)$ can be increased by an arbitrarily small amount while continuing to satisfy all of the constraints of the problem and leaving the policymaker strictly better off. If instead $c\left(\theta^N, z\right) > 0$, then $x\left(\theta^N, z\right)$ can be increased by an arbitrarily small amount while continuing to satisfy all of the constraints of the problem and leaving the policymaker strictly better off. However, if $V'\left(\theta^N, z\right) = \overline{V}$ and $c\left(\theta^N, z\right) = 0$, then part (i) of Lemma 6 implies that $c(\theta^n, z) = 0$ $\forall n$, which given that $C_{n+1,n} = 0 \ \forall n < N$ implies that $V'(\theta^n, z) = \overline{V} \ \forall n$. However, this contradicts the fact that $C_{N-1,N} > 0$. Therefore, $C_{N-1,N} = 0$.

Now suppose that $C_{\widetilde{n}-1,\widetilde{n}} = 0 \ \forall \widetilde{n} > n$ but that $C_{n-1,n} > 0$. Then this implies that $V'\left(\theta^{\widetilde{n}}, z\right) = \overline{V}$ and $c\left(\theta^{\widetilde{n}}, z\right) = 0 \ \forall \widetilde{n} \geq n$, and this follows by analogous reasoning as in the case for which $C_{N-1,N} > 0$. However, as before, part (i) of Lemma 6 implies that $c(\theta^n, z) = 0$ $\forall n$, which given that $C_{n+1,n} = 0 \ \forall n$ implies that $V'(\theta^n, z) = \overline{V} \ \forall n$, contradicting the fact that $C_{n-1,n} > 0$.

Proof of part (ii.a). Suppose that $i(z) = 0$ for some z. By Lemma 7, we can consider such a solution for which $C_{n+1,n} = 0 \ \forall n < N$. We establish that this is not possible in the below steps.

Step 1. It must be the case then that (19) is binding for some θ, since if this were not the case, one can perform the following perturbation to the solution α for the positive measure z for which $i(z) = 0$. Let $\widehat{i}(z, \epsilon) = \epsilon$ for some $\epsilon > 0$ arbitrarily small, and let $\widehat{c}(\theta^n, z, \epsilon) = c(\theta^n, \epsilon) + f(\epsilon) - \epsilon$, and leave the rest of the allocation unchanged. It can be easily verified that the perturbation satisfies $(15) - (21)$. Moreover, it makes households strictly better off. Therefore (19) must bind with equality for some θ, and by part (iv) of Lemma 6, it must bind for $\theta = \theta^1$.

Step 2. This implies that $x(\theta^n, z) = \theta^n \ \forall n$. To see why, consider a perturbation to the solution α which set $\widehat{i}(z, \epsilon) = \epsilon$ for some $\epsilon > 0$ arbitrarily small for all positive measure z for which $i(z) = 0$. Let $\widehat{x}(\theta^n, z, \epsilon)$ satisfy the following two equations

$$v(\widehat{x}(\theta^1, z, \epsilon)) - v(f(\widehat{i}(z, \epsilon) + \epsilon) + \theta^1) = v(x(\theta^1, z)) - v(f(i(z)) + \theta^1) \tag{39}$$

$$v(\widehat{x}(\theta^n, z, \epsilon)) - v(\widehat{x}(\theta^{n-1}, z, \epsilon) + \theta^n - \theta^{n-1}) = v(x(\theta^n, z)) - v(x(\theta^{n-1}, z) + \theta^n - \theta^{n-1}) \tag{40}$$

$\forall n > 1$. Finally, note that $\widehat{c}(\theta^n, z, \epsilon)$ is determined from the resource constraint

$$\widehat{c}(\theta^n, z, \epsilon) + \widehat{x}(\theta^n, z, \epsilon) = \omega - \widehat{i}(z, \epsilon) + f(\widehat{i}(z, \epsilon)) + \theta^n \ \forall n. \tag{41}$$

The rest of the allocation is left unchanged. It is straightforward to check that the perturbation satisfies $(16) - (21)$ so that it is sustainable and that it delivers a strictly higher continuation value to the incumbent. We can now show that it must make households strictly better off unless $x(\theta^n, z) = \theta^n \ \forall n$. Implicit differentiation of $(39) - (41)$ around $\epsilon = 0$ implies that

$$\frac{d\widehat{c}(\theta^n, z, 0)}{d\epsilon} = f'(0) - 1 - \frac{d\widehat{x}(\theta^n, z, 0)}{d\epsilon} \tag{42}$$

$$\frac{d\widehat{x}(\theta^1, z, 0)}{d\epsilon} = \frac{v'(\theta^1)}{v'(x(\theta^1, z))} f'(0) \leq f'(0), \text{ and} \tag{43}$$

$$\frac{d\widehat{x}(\theta^n, z, 0)}{d\epsilon} = \frac{v'\left(x(\theta^{n-1}, z) + \theta^n - \theta^{n-1}\right)}{v'(x(\theta^n, z))} \frac{d\widehat{x}(\theta^{n-1}, z, 0)}{d\epsilon} \leq f'(0) \ \forall n > 1. \tag{44}$$

The final inequality in (43) is a strict inequality if $\theta^1 > x\left(\theta^1, z\right)$, and this holds by the concavity of $v(\cdot)$. Analogous arguments imply that the final inequality in (44) is a strict inequality if $\theta^1 > x\left(\theta^1, z\right)$ or if $x\left(\theta^{\widetilde{n}}, z\right) < x\left(\theta^{\widetilde{n}-1}, z\right) + \theta^{\widetilde{n}} - \theta^{\widetilde{n}-1}$ for any $1 < \widetilde{n} \leq n$. From (17) and (42), it follows that the implied change in household consumption from an arbitrarily small increase in ϵ is positive if $x\left(\theta^n, z\right) < \theta^n$ for any n. Therefore, $x\left(\theta^n, z\right) = \theta^n \ \forall n$.

Step 3. It follows the fact that (19) is binding for some n and from the fact that $C_{n+1,n} = 0$ $\forall n < N$ that $V'\left(\theta^n, z\right) = V'\left(\theta^{n-1}, z\right) = \underline{V}$. However, one can show that this is suboptimal. From the proof of Lemma 5, it is clear that $V_0 > \underline{V}$. Consider then a perturbation to the solution α which set $\widehat{i}\left(z, \epsilon\right) = \epsilon$ for some $\epsilon > 0$ arbitrarily small for all positive measure z for which $i\left(z\right) = 0$. Moreover, let $\widehat{c}\left(\theta^n, z, \epsilon\right)$ is determined from

$$\widehat{c}\left(\theta^n, z, \epsilon\right) + x\left(\theta^n, z\right) = \omega - \widehat{i}\left(z, \epsilon\right) + f\left(\widehat{i}\left(z, \epsilon\right)\right) + \theta^n \ \forall n,$$

so that it is clear that consumption increases from the perturbation. Finally, let

$$\widehat{V}'\left(\theta^n, z, \epsilon\right) = \widehat{V}'(\theta^{n-1}, z, \epsilon) = (v(f(\widehat{i}\left(z, \epsilon\right)) + \theta^N) - v(\theta^N))/\beta + \underline{V}.$$

It can easily be verified that the perturbation satisfies (15) − (21). Moreover, for ϵ sufficiently low, the value of $J(\widehat{V}'\left(\theta^n, z, \epsilon\right)) - J\left(\underline{V}\right) = 0$, where this follows from part (ii) of Lemma 4. Therefore, the perturbation makes households strictly better off.

Proof of part (ii.b). Suppose that $c\left(\theta^n, z\right) = 0$ for some n. In order to rule out this case, we take the following approach. We establish that there exists a peturbed allocation which gives some welfare $V + \epsilon$ for $\epsilon \gtreqless 0$ small to the policymaker which makes households infinitely better off on the margin. This leads to a contradiction because if $\epsilon > 0$, this implies that $J\left(V\right)$ is upward sloping, violating Lemma 4, and if instead $\epsilon < 0$, this implies that $J\left(V\right)$ has a slope of $-\infty$, implying that $V = \overline{V}$, where this follows from the concavity of $J\left(V\right)$ established in Lemma 4. By Lemma 7, we can perturb around a solution for which $C_{n+1,n} = 0 \ \forall n < N$. By part (i) of Lemma 6 and (16), it follows that if $c\left(\theta^n, z\right) = 0$ for some n, then $c\left(\theta^1, z\right) = 0$ and there exists some n^* such that $c\left(\theta^n, z\right) = 0$ for all $n \leq n^*$. There are several cases to consider.

Case 1. Suppose that $x\left(\theta^n, z\right) > 0 \ \forall n$ and $n^* = N$ so that $c\left(\theta^n, z\right) = 0 \ \forall n$. This implies that $C_{n,n+1} = 0 \ \forall n < N$, so that $V'\left(\theta^{n+1}, z\right) = V'\left(\theta^n, z\right)$. Note that it cannot be that $V'\left(\theta^n, z\right) = \overline{V}$, since from part (i), this would imply that household welfare conditional on z is $u\left(0\right) + \beta u\left(\omega\right) / \left(1 - \beta\right)$, violating (20). It therefore follows given the concavity of $J\left(V\right)$ that the slope of $J\left(V'\left(\theta^n, z\right)\right)$ to the right of $V'\left(\theta^n, z\right)$ is well defined and bounded away from $-\infty$. Consider the following perturbation. Let $\widehat{c}\left(\theta^n, z, \epsilon\right) = \epsilon$ and $\widehat{x}\left(\theta^n, z, \epsilon\right) = x\left(\theta^n, z\right) - \epsilon$. Moreover, let

$$\widehat{V}'\left(\theta^n, z\right) = \sum_{n=1}^{N} \pi\left(\theta^n\right) \frac{v\left(x\left(\theta^n, z, \epsilon\right)\right) - v\left(\widehat{x}\left(\theta^n, z, \epsilon\right)\right)}{\beta} + V'\left(\theta^n, z\right).$$

Leave the rest of the allocation unchanged. It is straightforward to see that (16) − (21) is satisfied so that the perturbed allocation is sustainable. Moreover, it follows by Inada conditions and the fact that the slope of $J\left(V'\left(\theta^n, z\right)\right)$ to the right of $V'\left(\theta^n, z\right)$ is well defined so that by the Inada conditions on $u\left(\cdot\right)$ rate of increase in the welfare of the households is ∞ for arbitrarily small ϵ.

Case 2. Suppose that $x\left(\theta^n, z\right) > 0\ \forall n$ and $n^* < N$. Note that in this case, it must be that $V'\left(\theta^n, z\right) < \overline{V}$ for $n \leq n^*$, since if this were not the case, then $V'\left(\theta^n, z\right) = \overline{V}\ \forall n$ which given that $C_{n+1,n} = 0$ implies that $c\left(\theta^n, z\right) = 0\ \forall n$, leading to a contradiction. Consider the following perturbation. Let $\widehat{x}\left(\theta^n, z, \epsilon\right) = x\left(\theta^n, z\right) - \epsilon\ \forall n \leq n^*$. Moreover, for $n \leq n^*$, let

$$\widehat{V}'\left(\theta^n, z\right) = \frac{1}{\sum_{n=1}^{n^*} \pi\left(\theta^n\right)} \left(\sum_{n=1}^{n^*} \pi\left(\theta^n\right) \frac{v\left(x\left(\theta^n, z, \epsilon\right)\right) - v\left(\widehat{x}\left(\theta^n, z, \epsilon\right)\right)}{\beta} \right) + V'\left(\theta^n, z\right).$$

For $n = n^* + 1$, let $\widehat{x}\left(\theta^n, z, \epsilon\right)$ satisfy

$$\begin{aligned} v\left(\widehat{x}\left(\theta^n, z, \epsilon\right)\right) - v\left(\widehat{x}\left(\theta^{n-1}, z, \epsilon\right) + \theta^n - \theta^{n-1}\right) - \beta\widehat{V}'\left(\theta^{n-1}, z\right) = \\ v\left(x\left(\theta^n, z\right)\right) - v\left(x\left(\theta^{n-1}, z\right) + \theta^n - \theta^{n-1}\right) - \beta V'\left(\theta^{n-1}, z\right) \end{aligned}$$

and for $n > n^* + 1$, let $\widehat{x}\left(\theta^n, z, \epsilon\right)$ satisfy

$$v\left(\widehat{x}\left(\theta^n, z, \epsilon\right)\right) - v(\widehat{x}(\theta^{n-1}, z, \epsilon) + \theta^n - \theta^{n-1}) = v\left(x\left(\theta^n, z\right)\right) - v(x(\theta^{n-1}, z) + \theta^n - \theta^{n-1}).$$

Define $\widehat{c}\left(\theta^n, z, \epsilon\right) = c\left(\theta^n, z\right) - \left(\widehat{x}\left(\theta^n, z, \epsilon\right) - x\left(\theta^n, z\right)\right)$. Note that concavity implies that $\widehat{x}\left(\theta^n, z, \epsilon\right) > x\left(\theta^n, z\right)$ for $n > n^*$. Leave the rest of the allocation unchanged. It is straightforward to verify that (16) − (21) are satisfied so that the perturbed allocation is sustainable. The only issue to note in establishing this is that (17) cannot hold with equality for $n > n^*$ in the original allocation so that by continuity it is satisfied in the perturbed allocation. Note that if it were the case that (17) held with equality, then this would imply by part (iii) of Lemma 6 that (17) holds with equality for $n = 1$. But if that were true, $c\left(\theta^n, z\right) = 0$ for $n > n^*$ leading to a contradiction. Analogous arguments then as those of case 1 imply that the rate of increase in the welfare of the households is ∞ for arbitrarily small ϵ.

Case 3. Suppose that $x\left(\theta^n, z\right) = 0$ for some n. Note that if it is the case that $x\left(\theta^n, z\right) > 0$ for all n for which $c\left(\theta^n, z\right) = 0$, then the arguments case 2 can be utilized, since the same perturbation is feasible in this case. We are left then to consider the case for which $c\left(\theta^n, z\right) = x\left(\theta^n, z\right) = 0$ for some n. We can establish that this can only be true for $n = 1$. To see why, note that if $c\left(\theta^n, z\right) = 0$ for any n, then $c\left(\theta^1, z\right) = 0$ by part (i) of Lemma 6. Moreover, (16) implies that for all n for which $c\left(\theta^n, z\right) = 0$, $x\left(\theta^n, z\right)$ is a strictly increasing function of n. Therefore, we are left to consider the case for which $c\left(\theta^1, z\right) = x\left(\theta^1, z\right) = 0$ and $x\left(\theta^n, z\right) > 0$ if $c\left(\theta^n, z\right) = 0$ for $n > 1$. From (16), in order that $x\left(\theta^n, z\right) = 0$ and $c\left(\theta^n, z\right) = 0$, it must be that $i\left(z\right) > i^*$ so that $\omega - i\left(z\right) + f\left(i\left(z\right)\right) + \theta^1 = 0$. Now consider the following perturbation to α.

Let $\widehat{i}(z,\epsilon) = i(z) - \epsilon$ for $\epsilon > 0$ arbitrarily small and let

$$\widehat{c}(\theta^n, z, \epsilon) = c(\theta^n, z) + (f(\widehat{i}(z,\epsilon)) - \widehat{i}(z,\epsilon)) - (f(i(z)) - i(z)).$$

Leave the rest of the allocation as it is. It is straightforward to see that since (17) is a strict inequality $\forall n$ in the original allocation, $(15) - (21)$ are satisfied in the perturbed allocation by continuity for sufficiently small ϵ. Since $\widehat{c}(\theta^n, z, \epsilon) > c(\theta^n, z)$ it follows that household welfare is increased under the perturbation, implying that the original allocation was suboptimal.

Proof of part (ii.c). Suppose that (20) binds $\forall z$. This implies that $J(V) = u(\omega)/(1-\beta)$. Given that $J(V) \geq u(\omega)/(1-\beta)$ for all V, this implies given the fact that $J(V)$ is weakly concave and weakly decreasing that $J(V) = u(\omega)/(1-\beta)$ for all V. However, one can show that this is not true by constructing an allocation which provides households a welfare which strictly exceeds $u(\omega)/(1-\beta)$. Construct an equilibrium as in the proof of Lemma 1 with the exception that $c_t(q_t, z_t, \theta_t) = \omega + \epsilon$, and $x_t(q_t, z_t, \theta_t) = f(i^*) - i^* + \theta_t - \epsilon$ for all θ_t for some $\epsilon > 0$ sufficiently small. By the arguments in the proof of Lemma 1, the allocation satisfies all sustainability constraints. Moreover, it provides households with a welfare which strictly exceeds $u(\omega)/(1-\beta)$, violating the fact that $J(V) = u(\omega)/(1-\beta)$ for all V. ■

Lemma 9 *$J(V)$ is continuously differentiable in V for $V \in (\underline{V}, \overline{V})$.*

Proof. We use Lemma 1 of Benveniste and Scheinkman (1979) to prove the continuous differentiability of $J(V)$ for $V \geq V_0$. In particular we show that if there exists a function $Q(V+\epsilon)$ for $\epsilon \gtreqless 0$ which is differentiable, weakly concave, and satisfies

$$Q(V+\epsilon) \leq J(V+\epsilon) \tag{45}$$

for arbitrarily small values $|\epsilon|$ where (45) is an equality if $\epsilon = 0$. By Lemma 1 of Benveniste and Scheinkman (1979) then $J(V)$ is continuously differentiable at V. To do so we first characterize a potential solution α conditional on $V \in [V_0, \overline{V})$, construct a perturbed solution $\widehat{\alpha}(\epsilon)$, and then check that this perturbed solution satisfies $(16) - (21)$ for $\epsilon \gtreqless 0$ arbitrarily small and provides $V + \epsilon$ to the policymaker. From Lemmas 5, 7, and 8, we can perturb around an original solution α with the following properties for some positive measure z: $P(z) = 1$, $C_{n+1,n} = 0$ for all $n < N$, $i(z) > 0$, $c(\theta, z) > 0\ \forall\theta$, and (20) does not bind. We let $\widetilde{Z}$ correspond to such z and let $q_{\widetilde{Z}} = \Pr(z \in \widetilde{Z})$.

Case 1. Suppose that $x(\theta, z) > 0\ \forall\theta$ and $\forall z \in \widetilde{Z}$. Define $\widehat{\alpha}(\epsilon)$ as follows. If $z \notin \widetilde{Z}$, then the element of $\widehat{\alpha}(\epsilon)$ is identical to the element of α. If instead $z \in \widetilde{Z}$, then let $\widehat{\alpha}(\epsilon)$ be identical to α, with the exception that $\forall n$,

$$\begin{aligned}
\widehat{i}(z,\epsilon) &= i(z,\epsilon) + \xi^i(z,\epsilon), \\
\widehat{c}(\theta^n, z, \epsilon) &= c(\theta^n, z, \epsilon) + \xi^c(\theta^n, z, \epsilon), \text{ and} \\
\widehat{x}(\theta^n, z, \epsilon) &= x(\theta^n, z, \epsilon) + \xi^x(\theta^n, z, \epsilon)
\end{aligned}$$

for $\xi(z,\epsilon) = \left\{\xi^i(z,\epsilon), \{\xi^c(\theta^n,z,\epsilon), \xi^x(\theta^n,z,\epsilon), \}_{n=1}^N\right\}$ which satisfy

$$\sum_{n=1}^{N} \pi(\theta^n) v(x(\theta^n,z)) + \epsilon/q_{\widetilde{Z}} = \sum_{n=1}^{N} \pi(\theta^n) v(\widehat{x}(\theta^n,z,\epsilon)) \quad (46)$$

$$\widehat{c}(\theta^n,z,\epsilon) + \widehat{x}(\theta^n,z,\epsilon) = \omega - \widehat{i}(z,\epsilon) + f\left(\widehat{i}(z,\epsilon)\right) + \theta^n \ \forall n \quad (47)$$

$$v(\widehat{x}(\theta^{n+1},z,\epsilon)) - v(\widehat{x}(\theta^n,z,\epsilon) + \theta^{n+1} - \theta^n) = v(x(\theta^{n+1},z)) - v(x(\theta^n,z) + \theta^{n+1} - \theta^n), \forall n < N \quad (48)$$

$$v(\widehat{x}(\theta^1,z,\epsilon)) - v(f(\widehat{i}(z,\epsilon)) + \theta^1) = v(x(\theta^1,z)) - v(f(i(z)) + \theta^1) \quad (49)$$

for a given ϵ. (46) $-$ (49) corresponds to $2N+1$ equations and $2N+1$ unknowns, where the ξ's all equal 0 if ϵ equals 0. Since $u(\cdot)$ and $v(\cdot)$ are continuously differentiable, then every element of $\xi(z,\epsilon)$ for a given z is a continuously differentiable function of ϵ around $\epsilon = 0$. Note that it it straightforward to see from (46) $-$ (49) that that $\xi^i(z,\epsilon)$ and $\xi^x(\theta,z,\epsilon)$ rise in ϵ. The change in $\xi^c(\theta,z,\epsilon)$ is ambiguous. Define $Q(V+\epsilon)$ as the household welfare implied by the perturbed allocation $\widehat{\alpha}(\epsilon)$. It follows that $Q(V) = J(V)$ and that $Q(V+\epsilon)$ is continuously differentiable around $\epsilon = 0$. Note that because $J(V)$ is concave with $Q(V) = J(V)$, satisfaction of (45) implies that $Q(V)$ is also locally concave around V.

We are left to verify that (45) is satisfied. In order to do this, verify that every element of $\widehat{\alpha}(\epsilon)$ satisfies (16) $-$ (21) for a given promised value $V+\epsilon$ since this implies that $\widehat{\alpha}(\epsilon)$ is a potential sustainable solution to the program so that (45) must hold. Satisfaction of (15) under α together with (46) implies that $\widehat{\alpha}(\epsilon)$ satisfies (15) for continuation value $V+\epsilon$. Satisfaction of (47) implies that $\widehat{\alpha}(\epsilon)$ satisfies (16). Now let us check that (18) is satisfied. To do this, we appeal to part (ii) of Lemma 6 and simply check that $C_{n+1,n} \geq 0$ and $C_{n,n+1} \geq 0$ under $\widehat{\alpha}(\epsilon)$. Given that $C_{n+1,n} = 0$ under α, (48) guarantees that $C_{n+1,n} = 0$ under $\widehat{\alpha}(\epsilon)$. Note furthermore that if $C_{n,n+1} > 0$ under α, then $C_{n,n+1} > 0$ under $\widehat{\alpha}(\epsilon)$ for sufficiently small ϵ by continuity. We are left to consider the situation for which $C_{n,n+1} = 0$ under α. In this case, $x\left(\theta^{n+1},z\right) = x(\theta^n,z) + \theta^{n+1} - \theta^n$, which from (48) means that $\widehat{x}\left(\theta^{n+1},z,\epsilon\right) = \widehat{x}(\theta^n,z,\epsilon) + \theta^{n+1} - \theta^n$ so that $C_{n,n+1} = 0$ under $\widehat{\alpha}(\epsilon)$ as well. Thus, (18) is satisfied under $\widehat{\alpha}(\epsilon)$. Now let us check that (17) is satisfied under $\widehat{\alpha}(\epsilon)$. By part (iii) of Lemma 6, it is sufficient to check that this is the case if $\theta = \theta^1$ since (18) is satisfied. This is guaranteed by the fact that (17) is satisfied under α and by (49). Now let us check (19). By part (iv) of Lemma 6, it is sufficient to check that this is the case if $\theta = \theta^1$ since (17) and (18) are satisfied. This is also guaranteed by the fact that (19) is satisfied under α and by (49). Since (20) is slack under α, then it is also slack under $\widehat{\alpha}(\epsilon)$ for sufficiently small ϵ. Finally since (21) is satisfied under α, it is also satisfied under $\widehat{\alpha}(\epsilon)$. Since the perturbed allocation satisfies all of the constraints, it follows that (45) holds.

Case 2. Suppose now that it is the case that $x(\theta,z) = 0$ for some positive measure z, and with some abuse of notation, relabel $\widetilde{Z}$ as the set of all such z's. We will prove that in this case, $J'(V) = 0$. Define $\widehat{\alpha}(\epsilon)$ for $\epsilon > 0$ analogously to case 1, where this is feasible since

$\widehat{x}(\theta^n, z, \epsilon) > x(\theta^n, z)\ \forall n$. It follows by implicit differentiation given the Inada conditions on $v(\cdot)$ and the fact that $\widehat{i}(z, \epsilon)$ and $\widehat{c}(\theta^n, z, \epsilon)$ are interior that

$$\lim_{\epsilon > 0, \epsilon \to 0} \frac{Q(V + \epsilon) - Q(V)}{\epsilon} = 0 \leq \lim_{\epsilon > 0, \epsilon \to 0} \frac{J(V + \epsilon) - J(V)}{\epsilon}, \tag{50}$$

where we have used the fact that $Q(V + \epsilon) \leq J(V + \epsilon)$ for $\epsilon > 0$ and $Q(V) = J(V)$. Given that $J(V)$ is weakly decreasing, it follows that the last weak inequality in (50) binds with equality. Since $J(V)$ is weakly decreasing and weakly concave, it follows that

$$0 \geq \lim_{\epsilon > 0, \epsilon \to 0} \frac{J(V) - J(V - \epsilon)}{\epsilon} \geq \lim_{\epsilon > 0, \epsilon \to 0} \frac{J(V + \epsilon) - J(V)}{\epsilon} = 0,$$

which implies that $J(V)$ is differentiable at V with $J'(V) = 0$. ■

C Proofs of Section 4

C.1 Proof of Lemma 2

We consider the solution to (14) which ignores constraint (18) so that there is no private information. We first prove the following preliminary result.

Lemma 10 *The solution to (14) which ignores constraint (18) has the following features:*

1. *If $V = \overline{V}$, then $c(\theta^n, z) = c^{\max}(\theta^n)$, $x(\theta^n, z) = x^{\max}(\theta^n)$, $i(z) = i^*$, and $V'(\theta^n, z) = \overline{V}\ \forall z$,*
2. *If $V < \overline{V}$, but (20) binds conditional on z, then $c(\theta^n, z) = c^{\max}(\theta^n)$, $x(\theta^n, z) = x^{\max}(\theta^n)$, $i(z) = i^*$, and $V'(\theta^n, z) = \overline{V}$.*

Proof. Proof of part (i). Let

$$\overline{V}^* = \max_{\{c_t, x_t, i_t\}_{t=0}^{\infty}} E_0 \left\{ \sum_{t=0}^{\infty} \beta^t v(x_t) \right\}$$

$$\text{s.t.}$$

$$c_t + x_t = \omega - i_t + f(i_t) + \theta_t\ \forall t,$$

$$\text{and } E_t \left\{ \sum_{k=0}^{\infty} \beta^k u(c_{t+k}) \right\} \geq u(\omega) / (1 - \beta)\ \ \forall t.$$

If it is the case that $\overline{V} = \overline{V}^*$, then starting from $V = \overline{V}$, the unique solution to (14) which ignores constraint (18) coincides with the unique solution to the above program. Note that such a solution satisfies the features described in the statement of part (i) of the lemma.

We now verify that $\overline{V} = \overline{V}^*$. To do this, we consider the allocation described in the statement of the lemma and verify that it satisfies all of the sustainability constraints. Note that since $c^{\max}(\theta^n) > 0$, (17) is satisfied $\forall n$, and this follows from Assumption 3. (16) and (20) are also clearly satisfied, so we are left to consider the limited commitment constraint (19), and this is guaranteed by Assumption 4. Therefore, the solution satisfies sustainability constraints so that $\overline{V} = \overline{V}^*$.

Proof of part (ii). Now suppose that given V, (20) binds conditional on z so that households receive a continuation welfare of $u(\omega)/(1-\beta)$. Optimality then requires that the policymaker receive a continuation value of $\overline{V}$ conditional on z, since otherwise it would be possible to make the policymaker strictly better off while leaving the households as well off, violating the fact that $J(V)$ is downward sloping. By part (i), there is a unique method of providing this continuation value described in the statement of the lemma. ■

We now proceed in proving Lemma 2.

Proof of part (i) Let λ, $\pi(\theta^n)\nu(\theta^n,z)\,dz$, $\pi(\theta^n)\kappa(\theta^n,z)\,dz$, $\pi(\theta^n)\psi(\theta^n,z)\,dz$, and $\eta(z)\,dz$ correspond to the Lagrange multipliers on constraints (15), (16), (17), (19), and (20), respectively. Let $\beta\pi(\theta^n)\underline{\mu}(\theta^n,z)\,dz$, $\beta\pi(\theta^n)\overline{\mu}(\theta^n,z)\,dz$, and $\pi(\theta^n)\zeta(\theta^n,z)\,dz$ correspond to the Lagrange multipliers on the constraints that $V'(\theta^n,z) \geq \underline{V}$, $V'(\theta^n,z) \leq \overline{V}$, and $x(\theta^n,z) \geq 0$, respectively. Analogous arguments as those of Lemma 8 imply that the constraint that $i(z) \geq 0$ and $c(\theta,z) \geq 0$ is always slack and can therefore be ignored. First order conditions imply:

$$u'(c(\theta^n,z))(P(z)+\eta(z)) = \nu(\theta^n,z), \tag{51}$$

$$(P(z)\lambda + \psi(\theta^n,z))\,v'(x(\theta^n,z)) = \nu(\theta^n,z) + \kappa(\theta^n,z) - \zeta(\theta^n,z) \tag{52}$$

$$f'(i(z)) - 1 = \frac{\sum_{n=1}^{N} \pi(\theta^n)\left(-\kappa(\theta^n,z) f'(i(z)) + \psi(\theta^n,z) v'(f(i(z))+\theta^n) f'(i(z))\right)}{\sum_{n=1}^{N} \pi(\theta)\nu(\theta^n,z)} \tag{53}$$

$$J'\left(V'(\theta^n,z)\right)(P(z)+\eta(z)) = \left\{-P(z)\lambda - \psi(\theta^n,z) - \underline{\mu}(\theta^n,z) + \overline{\mu}(\theta^n,z)\right\} \tag{54}$$

and the Envelope condition yields:

$$J'(V) = -\lambda. \tag{55}$$

To prove that a distortion emerges, we consider the situation with $V = V_0$ which occurs at $t = 0$. Analogous argument as those of part (ii) of Lemma 4 imply that $J'(V_0) = 0$, hence at V_0, $\lambda = 0$. Moreover, analogous arguments as those of Lemma 5 imply that $P(z) = 1$ $\forall z$ if $V \geq V_0$. Finally, note that if $V = V_0$, then necessarily $\eta(z) = 0$. This is because if this were not the case, then (20) binds for some z and households would be receiving a continuation welfare of $u(\omega)/(1-\beta) < \overline{J}$ with positive probability, violating the fact that $J(V_0) = \overline{J}$ by definition. Now suppose by contradiction that $i(z) \geq i^*$ starting for $V = V_0$. We establish that this is not possible in a two steps.

Step 1. It is not possible that $\psi(\theta^n,z) = 0$ $\forall n$. Suppose this were the case. Then (52) and (53) given that $\lambda = 0$ imply that $\zeta(\theta^n,z) > 0$ so that $x(\theta^n,z) = 0$ $\forall n$. Moreover, (54) given that $\lambda = \eta(z) = 0$ implies that $J'(V'(\theta^n, z)) = 0$.[40] Since V_0 corresponds to the highest continuation value V which the policymaker can receive subject to $J'(V) = 0$, it follows then that for all θ^n and z: $V'(\theta^n, z) \leq V_0$, so that $V_0 \leq v(0) / (1 - \beta)$. However, if this is the case, together with Assumption 1 it implies that (19) is violated, leading to a contradiction.

Step 2. Consider some n for $\psi(\theta^n, z) > 0$ so that (19) binds for such n. The same arguments as those of part (v) of Lemma 6 imply that constraint (17) can be ignored in this case so that $\kappa(\theta^n, z) = 0$ for such n. Moreover, note that for all other n for which $\psi(\theta^n, z) = 0$, $x(\theta^n, z) = 0$ by the same reasoning as in step 1, and it trivially follows that (17) cannot bind for such n either. Therefore, $\kappa(\theta^n, z) = 0$ $\forall n$. However, if this is the case, then (53) implies that $i(z) < i^*$, leading to a contradiction.

Proof of part (ii) To prove this second result, we first establish that $J'(V'(\theta^n, z)) \leq J'(V)$ $\forall z$ and $\forall n$. Suppose first that $\eta(z) > 0$. This implies given Lemma 10 that $V'(\theta^n, z) = \overline{V}$, and from the concavity of $J(V)$, this implies that $J'(V'(\theta^n, z)) \leq J'(V)$. If instead $\eta(z) = 0$, then (54) and the Envelope condition imply that if it is not the case that $V'(\theta^n, z) = \overline{V}$, then

$$J'\left(V'(\theta^n, z)\right) = J'(V) - \psi(\theta^n, z) \leq J'(V).$$

Therefore, given a stochastic sequence $\{V_t\}_{t=0}^{\infty}$, it must be that the associated stochastic sequence $\{J'(V_t)\}_{t=0}^{\infty}$ is monotonical decreasing. As such, from (55), this means that the sequence $\{\lambda_t\}_{t=0}^{\infty}$ is itself either converging or declining towards $-\infty$. This leaves us with two cases to consider.

Case 1. Suppose that λ_t converges to a finite number. From (54) and (55), this implies that $\{\psi_t\}_{t=0}^{\infty}$ is a sequence which converges to zero. Note that since V_t is bounded, there exists a convergent subsequence of $\{V_t\}_{t=0}^{\infty}$ which converges to some limiting continuation value V. By continuity, there is some limiting associated Lagrange multiplier $\psi = 0$. Given V, the first order condition for investment (53) implies that $i(z) \geq i^*$ in this limiting allocation. If this is the case, then given Assumption 3, it follows that (17) cannot bind in this limiting allocation, and therefore from (53), $i(z) = i^*$ in the limiting allocation.

Case 2. Suppose that λ_t diverges towards ∞ so that therefore $J'(V_t)$ diverges towards $-\infty$. In this case, it must be that $\{V_t\}_{t=0}^{\infty}$ converges towards $\overline{V}$, and this follows from the concavity of $J(\cdot)$. However, by the reasoning of Lemma 10, the limiting allocation here also entails $i(z) = i^*$.■

C.2 Proof of Lemma 3

We consider the solution to (14) which ignores constraint (19). Let λ, $\pi(\theta^n)\,\nu(\theta^n, z)\,dz$, $\pi(\theta^n)\,\kappa(\theta^n, z)\,dz$, and $\eta(z)\,dz$ correspond to the Lagrange multipliers on constraints (15),

[40] This follows from the fact that since $J'(V'(\theta^n, z))$ is non-positive, $\overline{\mu}(\theta^n, z) = 0$, and moreover, it is not possible that $\underline{\mu}(\theta^n, z) > 0$, since from (54), this would lead to $J'(V'(\theta^n, z)) < 0$ and thus $V'(\theta^n, z) > \underline{V}$, which is a contradiction with $\underline{\mu}(\theta^n, z) > 0$.

(16), (17), and (20), respectively. Let $\beta\pi(\theta^n)\underline{\mu}(\theta^n,z)\,dz$, $\beta\pi(\theta^n)\overline{\mu}(\theta^n,z)\,dz$, and $\pi(\theta^n)\zeta(\theta^n,z)\,dz$ correspond to the Lagrange multipliers on the constraints that $V'(\theta^n,z)\geq\underline{V}$, $V'(\theta^n,z)\leq\overline{V}$, and $x(\theta^n,z)\geq 0$, respectively. Analogous arguments as those of Lemma 8 imply that the constraint that $i(z)\geq 0$ and $c(\theta,z)\geq 0$ is always slack and can therefore be ignored. Finally, by part (ii) of Lemma 6, we need only consider the local constraints for (18). Let $\pi\left(\theta^{n+1}\right)\phi\left(\theta^{n+1},\theta^n,z\right)dz$ and $\pi(\theta^n)\phi\left(\theta^n,\theta^{n+1},z\right)dz$ correspond to the Lagrange multipliers on the downward and upward incentive compatibility constraint, respectively, where we define $\phi\left(\theta^n,\theta^{n-1},z\right)=0$ if $n=1$ and $\phi\left(\theta^{n+1},\theta^n,z\right)=0$ if $n=N$. The first order condition with respect to $i(z)$ is:

$$f'(i(z))-1=\frac{\sum_{n=1}^{N}\pi(\theta^n)\left(-\kappa(\theta^n,z)f'(i(z))\right)}{\sum_{n=1}^{N}\pi(\theta)\nu(\theta^n,z)}.$$

Therefore, if it is the case that $i(z)\neq i^*$, this implies that $i(z)>i^*$ and that (17) binds for some θ^n. However, by Assumption 3, if this is the case, then this violates the non-negativity of consumption, leading to a contradiction. Therefore, $i(z)=i^*\ \forall z$.■

D Proofs of Section 5

D.1 Proof of Proposition 2

Proof of part (i) To show that distortions emerge along the equilibrium path, we pursue the same strategy as in the proof of part (i) of Lemma 2. Define all of the Lagrange multipliers as in the proofs of Lemma 2 and Lemma 3. First order conditions yield:

$$u'(c(\theta^n,z))\left(P(z)+\eta(z)\right)=\nu(\theta^n,z), \tag{56}$$

$$\left\{\begin{array}{c}\left(\begin{array}{c}P(z)\lambda+\phi\left(\theta^n,\theta^{n+1},z\right)\\+\phi\left(\theta^n,\theta^{n-1},z\right)+\psi(\theta^n,z)\end{array}\right)v'(x(\theta^n,z))\\-\phi\left(\theta^{n-1},\theta^n,z\right)v'\left(x(\theta^n,z)+\theta^{n-1}-\theta^n\right)\frac{\pi(\theta^{n-1})}{\pi(\theta^n)}\\-\phi\left(\theta^{n+1},\theta^n,z\right)v'\left(x(\theta^n,z)+\theta^{n+1}-\theta^n\right)\frac{\pi(\theta^{n+1})}{\pi(\theta^n)}\end{array}\right\}=\nu(\theta^n,z)+\kappa(\theta^n,z)-\zeta(\theta^n,z) \tag{57}$$

$$f'(i(z)) - 1 = \frac{\sum_{n=1}^{N} \pi(\theta^n)\left(-\kappa(\theta^n, z) f'(i(z)) + \psi(\theta^n, z) v'(f(i(z)) + \theta^n) f'(i(z))\right)}{\sum_{n=1}^{N} \pi(\theta) \nu(\theta^n, z)} \tag{58}$$

$$J'\left(V'(\theta^n, z)\right)(P(z) + \eta(z)) = \left\{ \begin{array}{c} -P(z)\lambda - \phi\left(\theta^n, \theta^{n-1}, z\right) - \phi\left(\theta^n, \theta^{n+1}, z\right) \\ +\phi\left(\theta^{n-1}, \theta^n, z\right) \frac{\pi(\theta^{n-1})}{\pi(\theta^n)} + \phi\left(\theta^{n+1}, \theta^n, z\right) \frac{\pi(\theta^{n+1})}{\pi(\theta^n)} \\ -\psi(\theta^n, z) - \underline{\mu}(\theta^n, z) + \overline{\mu}(\theta^n, z) \end{array} \right\} \tag{59}$$

and the Envelope condition yields:

$$J'(V) = -\lambda. \tag{60}$$

The following lemma, which is implied by Assumption 3, is useful for the remainder of our analysis.

Lemma 11 *Suppose that $i(z) \neq i^*$. Then (19) holds as an equality for $\theta^n = \theta^1$ with $i(z) < i^*$.*

Proof. Suppose that $i(z) \neq i^*$ and suppose first that $i(z) > i^*$. Given (58), this is only possible if (17) holds with equality. However, if this is the case, Assumption 3 together with (16) then imply that $c(\theta^n, z) < 0$, leading to a contradiction. Therefore, it must be that if $i(z) \neq i^*$, then $i(z) < i^*$. Now suppose that $i(z) < i^*$ but that (19) holds as a strict inequality for $\theta^n = \theta^1$. Then from part (iv) of Lemma 6, (19) is a strict inequality for all θ^n, which from (58) implies that $i(z) \geq i^*$, which is a contradiction. This establishes that (19) must hold with equality and that $i(z) < i^*$. ■

Now to prove our result, note as a reminder that from part (ii) of Lemma 4, it is the case that that $J'(V_0) = 0$. Analogous arguments to those in the proof of Lemma 2 imply that $\eta(z) = 0$. Now suppose that at V_0, $i(z) \geq i^*$, where from Lemma 11, this is only possible if $i(z) = i^*$. To show this is not the case, we proceed in two analogous steps to those in the proof of Lemma 2.

Step 1. It is not possible that $\psi(\theta^n, z) = 0\ \forall n$. Suppose this was the case. From (59) and (60), this implies that given z,

$$\sum_{n=1}^{N} \pi(\theta^n) J'\left(V'(\theta^n, z)\right) = J'(V_0) - \sum_{n=1}^{N} \pi(\theta^n) \overline{\mu}(\theta^n, z) + \sum_{n=1}^{N} \pi(\theta^n) \overline{\mu}(\theta^n, z). \tag{61}$$

Given that $J(\cdot)$ is weakly decreasing it follows that the left hand side of equation (61) is weakly negative. This implies that $\forall n$, $\overline{\mu}(\theta^n, z) = 0$ and $J'\left(V'(\theta^n, z)\right) = 0$. From part (ii) of Lemma 4, this implies that $V'(\theta^n, z) \leq V_0\ \forall\ n$. Therefore, from (59) for $n = 1$ it must be that $\phi\left(\theta^1, \theta^2, z\right) = \phi\left(\theta^2, \theta^1, z\right) \frac{\pi(\theta^2)}{\pi(\theta^1)}$. This implies from (59) $\forall n$ that $\phi\left(\theta^n, \theta^{n+1}, z\right) = \phi\left(\theta^{n+1}, \theta^n, z\right) \frac{\pi(\theta^{n+1})}{\pi(\theta^n)}$ for all $n < N$. We now show that $\phi\left(\theta^n, \theta^{n-1}, z\right) = 0$ for all $n > 1$. Suppose not, then $\phi\left(\theta^N, \theta^{N-1}, z\right) > 0$. Since it must also be the case then that $\phi\left(\theta^{N-1}, \theta^N, z\right) > 0$, this means that

$$x(\theta^{N-1}) = x(\theta^N) - (\theta^N - \theta^{N-1}) < x(\theta^N). \tag{62}$$

Now consider (57) for $n = N$ given that $\phi\left(\theta^N, \theta^{N-1}, z\right) > 0$:

$$0 > \phi(\theta^N, \theta^{N-1}, z)[v'(x(\theta^N, z)) - v'(x(\theta^N, z) - (\theta^N - \theta^{N-1}))] = \nu(\theta^N, z) + \kappa(\theta^N, z) - \zeta(\theta^N, z). \tag{63}$$

From (56), it must be that $\nu(\theta^N, z) > 0$ and constraint (17) implies that $\kappa(\theta^N, z) \geq 0$, which means that for (63) to hold, it must be that $\zeta(\theta^N, z) > 0$ so that $x(\theta^N, z) = 0$. However, this implies from (62) that $x(\theta^{N-1}) < 0$, which is not possible. Therefore, $\phi(\theta^N, \theta^{N-1}, z) = 0$. Now suppose that $\phi(\theta^{N-1}, \theta^{N-2}, z) > 0$. Analogous reasoning to the above implies analogous conditions to (62) and (63) for $N - 1$. But this leads to a contradiction since it implies that $x(\theta^{N-2}) < 0$, which is not possible. Similar arguments imply that $\phi\left(\theta^n, \theta^{n-1}, z\right) = 0$ for all $n > 1$.

Now consider (57) given that $\phi\left(\theta^n, \theta^{n-1}, z\right) = \phi\left(\theta^{n-1}, \theta^n, z\right) = 0$ for all $n > 1$. Since $\nu\left(\theta^n, z\right) > 0$ and $\kappa\left(\theta^n, z\right) \geq 0$, this means that $\zeta\left(\theta^n, z\right) > 0$ and $x\left(\theta^n, z\right) = 0$ for all n. Therefore, conditional on z and θ^n, the policymaker receives a continuation

$$v\left(0\right) + \beta V'\left(\theta^n, z\right) \leq v\left(0\right) / \left(1 - \beta\right).$$

However, if that were true, then (19) is violated by Assumption 1.

Step 2. We have established that $\psi\left(\theta^n, z\right) > 0$ for some θ^n. From part (iv) of Lemma 6, this would only be the case for $n = 1$, so that $\psi\left(\theta^1, z\right) > 0$ and $\psi\left(\theta^n, z\right) = 0$ $\forall n > 1$. Moreover, from parts (iii) and (v) of Lemma 6, constraint (17) is made redundant for all n and can be ignored since (19) binds with an equality, so that $\kappa(\theta^n, z) = 0$ $\forall n$. Given that $\nu\left(\theta^n, z\right) > 0$ from (56), this means that the right hand side of (58) is positive so that $i\left(z\right) < i^*$.■

Proof of part (ii) Suppose it were the case that $\lim_{t\to\infty} \Pr\left\{i_{t+k} = i^* \ \forall k\right\} > 0$. This would imply that there exist a stochastic sequence $\{i_t\}_{t=0}^{\infty}$ with positive measure which converges to i^*. Associated with such a stochastic sequence is a stochastic sequence $\{V_t\}_{t=0}^{\infty}$ which must includes within it at least one convergent stochastic subsequence, where this follows from the fact that V_t is bounded. Let Ψ correspond to the entire set of all limiting values of convergent stochastic subsequences of $\{V_t\}_{t=0}^{\infty}$ for every single stochastic sequence $\{i_t\}_{t=0}^{\infty}$ which converges to i^*. It follows by the continuity of the policy function that the solution to (14) – (21) starting from some $V \in \Psi$ yields a solution which admits $i\left(z\right) = i^*$ $\forall z$. Moreover, it must be that $V'\left(\theta, z\right) \in \Psi$ $\forall \theta$ and $\forall z$ with positive measure since $V'\left(\theta, z\right)$ must itself be within the set of limits of stochastic subsequences. Let $\underline{V}'$ correspond to the infinum of continuation values in Ψ. It is clear that $\underline{V}' > V_0$, since if this were not the case, this would imply by the proof of part (i) given that $\lambda = 0$ by part (ii) of Lemma 4 that $i\left(z\right) \neq i^*$ for some $V \in \Psi$, leading to a contradiction. The strategy of our proof is to show that $\underline{V}'$ does not exist so as to create a contradiction. Before proceeding, we prove the following preliminary lemma.

Lemma 12 *Consider the solution to (14) – (21) given $V \in \left(V_0, \overline{V}\right)$. Suppose that in the solution, (19) is slack $\forall \theta$ and $\forall z$. Then it must be that the solution admits $J'\left(V'\left(\theta^1, z\right)\right) > J'\left(V\right)$ $\forall z$.*

Proof. Note that by Lemma 5, the solution must admit $P(z) = 1\ \forall z$. There are two cases to consider.

Case 1. Suppose that the elements of α do not vary with z. Now suppose that $J'\left(V'\left(\theta^1, z\right)\right) \leq J'(V)\ \forall z$. Part (i) of Lemma 6 together with the weak concavity of $J(V)$ then implies that $J'\left(V'\left(\theta^n, z\right)\right) \leq J'(V)\ \forall n$ and $\forall z$, where we have used the fact that α does not depend on z. From Lemma 7, one can perturb such a solution without changing households' welfare and continuing to satisfy the constraints of the problem by changing the values of $V'\left(\theta^n, z\right)$ so that $C_{n+1,n} = 0$ for all $n < N$. Note that this perturbation weakly increases the value of $V'\left(\theta^1, z\right)$ so that it remains the case that $J'\left(V'\left(\theta^1, z\right)\right) \leq J'(V)\ \forall z$.

Now consider such a solution. Such a solution corresponds to the solution to the following problem, where λ corresponds to the Lagrange multiplier on constraint (15):

$$J(V) = \max_{\alpha} \left\{ \int_0^1 \left[\begin{array}{c} \sum_{n=1}^N \pi\left(\theta^n\right)\left(u\left(c\left(\theta^n, z\right)\right) + \beta J\left(V'\left(\theta^n, z\right)\right)\right) \\ +\lambda \left(\sum_{n=1}^N \pi\left(\theta^n\right)\left(v\left(x\left(\theta^n, z\right)\right) + \beta V'\left(\theta^n, z\right)\right)\right) \end{array} \right] dz \right\}$$

s.t.

$$c\left(\theta^n, z\right) + x\left(\theta^n, z\right) = \omega - i(z) + f(i(z)) + \theta^n\ \forall n, z, \tag{64}$$

$$x\left(\theta^n, z\right) \leq f(i(z)) + \theta^n\ \forall n, z, \tag{65}$$

$$x(\theta^{n+1}, z) \leq x\left(\theta^n, z\right) + \theta^{n+1} - \theta^n\ \forall n < N, z, \tag{66}$$

$$v(x(\theta^{n+1}, z)) + \beta V'(\theta^{n+1}, z) = v(x(\theta^n, z) + \theta^{n+1} - \theta^n) + \beta V'\left(\theta^n, z\right)\ \ \forall n < N, z, \tag{67}$$

$$\sum_{n=1}^N \pi\left(\theta^n\right) u\left(c\left(\theta^n, z\right) + \beta J\left(V'\left(\theta^n, z\right)\right)\right) \geq u(\omega) / (1-\beta)\ \ \forall z, \tag{68}$$

$$\text{and } V'\left(\theta^n, z\right) \in \left[\underline{V}, \overline{V}\right]\ \ \forall n, z. \tag{69}$$

The above program differs from the general program in the following fashion: It takes into account that replacement never occurs; it has removed constraints which do not bind; it has substituted (15) into the objective function taking into account that λ is the Lagrange multiplier on (15); and it has replaced constraint (18) with constraints (66) and (67) by using part (ii) of Lemma 6 and Lemma 7.

Since $V > V_0$, we only need consider the case for which $\lambda > 0$. This is because if $\lambda = 0$, then $V \leq V_0$, where this follows from the Envelope condition in (60) and part (ii) of Lemma 4. Define Lagrange multipliers analogously as in step 1. First order conditions with respect to $V'\left(\theta^n, z\right)$ yield:

$$J'\left(V'\left(\theta^n, z\right)\right)(1 + \eta(z)) = -\lambda - \phi\left(\theta^n, \theta^{n-1}, z\right) + \phi\left(\theta^{n+1}, \theta^n, z\right) + \overline{\mu}\left(\theta^n, z\right), \tag{70}$$

where we have taken into account that the fact that $V'\left(\theta^n, z\right) \geq V$ implies that $V'\left(\theta^n, z\right) > \underline{V}$ so that $\underline{\mu}\left(\theta^n, z\right) = 0$. Since the allocation does not depend on z, it is the case that $\eta(z) = 0$, since otherwise (68) binds $\forall z$ so that $J(V) = u(\omega)/(1-\beta)$. The Envelope condition yields: $J'(V) = -\lambda$. From (70), since $J'(V'(\theta^n, z)) \leq J'(V)$ for $n = 1$, this implies that $\phi\left(\theta^2, \theta^1, z\right) \leq 0$. For $n = 2$, this implies that $\phi\left(\theta^3, \theta^2, z\right) \leq \phi\left(\theta^2, \theta^1, z\right) \leq 0$, and forward induction implies that $\phi\left(\theta^N, \theta^{N-1}, z\right) \leq \phi\left(\theta^2, \theta^1, z\right) \leq 0$. (70) for $n = N$ given that $J'(V'(\theta^n, z)) \leq J'(V)$ requires $\phi\left(\theta^N, \theta^{N-1}, z\right) \geq 0$, which thus implies that $\phi\left(\theta^{n+1}, \theta^n, z\right) = 0 \ \forall n < N$. Therefore, constraint (67) can be ignored.

Let us assume and later verify that constraint (66) can also be ignored if constraint (67) can be ignored. First order conditions with respect to $c(\theta^n, z)$ and $x(\theta^n, z)$ together with (64) imply that

$$\lambda v'(x(\theta^n, z)) \geq u'(\omega - i(z) + f(i(z)) + \theta^n - x(\theta^n, z)), \tag{71}$$

which is a strict inequality only if $x(\theta^n, z) = f(i(z)) + \theta^n$. It is easy to verify that constraint (66) is satisfied under such a solution. This is because the value of $x(\theta^n, z)$ for which (71) binds is such that $x(\theta^n, z) - \theta^n$ is strictly declining in θ^n, where this follows from the concavity of $v(\cdot)$ and $u(\cdot)$. Consequently, if (71) is a strict inequality with $x(\theta^n, z) = f(i(z)) + \theta^n$ for some n, then it follows that $x\left(\theta^{n-k}, z\right) = f(i(z)) + \theta^{n-k}$ for all $k < n-1$. It therefore follows that there exists some n^* (which could be 1 or N) such that $x(\theta^n, z) = f(i(z)) + \theta^n$ if $n < n^*$ and $x(\theta^n, z) < f(i(z)) + \theta^n$ if $n \geq n^*$ with $x(\theta^n, z) - \theta^n$ is strictly declining in θ^n if $n \geq n^*$. Therefore, this means that (66) is satisfied under this solution.

Note that given the strict concavity of the program and the convexity of the constraint set with respect to $c(\theta^n, z)$, $x(\theta^n, z)$, and $i(z)$ since (67) is ignored, it follows that these values are uniquely defined conditional on λ. Moreover, since $J'(V'(\theta^n, z)) = -\lambda \ \forall n$, it follows by forward iteration on the recursive program that the same $c(\theta^n, z)$, $x(\theta^n, z)$, and $i(z)$ are chosen at all future dates. Therefore, $V'(\theta^n, z) = V$. Given that $C_{n+1,n} = 0$ under this solution, this can only be true if $x\left(\theta^{n+1}, z\right) = x(\theta^n, z) + \theta^{n+1} - \theta^n \ \forall n < N$, which given the above reasoning is only true if $x(\theta^n, z) = f(i(z)) + \theta^n \ \forall n$. If that is the case, then the welfare of households given V equals $u(\omega - i(z))/(1-\beta)$, which means that for (68) to be satisfied, it must be the case that $i(z) = 0$, but this violates part (ii) of Lemma 8.

Case 2. Now suppose that α varies with the value of z. We can show first that conditional on z, it must be that

$$J'\left(V'\left(\theta^1, z\right)\right) > J'\left(\sum_{n=1}^{N} \pi(\theta^n)\left(v(x(\theta^n, z)) + \beta V'(\theta^n, z)\right)\right). \tag{72}$$

To see why, note that if such an allocation is instead chosen for all z, it provides a continuation value

$$V = \sum_{n=1}^{N} \pi(\theta^n)\left(v(x(\theta^n, z)) + \beta V'(\theta^n, z)\right) \tag{73}$$

to the policymaker and optimality requires that it provides social welfare equal to

$$J\left(\sum_{n=1}^{N}\pi\left(\theta^{n}\right)\left(v\left(x\left(\theta^{n},z\right)\right)+\beta V'\left(\theta^{n},z\right)\right)\right). \tag{74}$$

This is because exceeding this social welfare is not possible by definition and if the implied social welfare were below the above value, the original solution would be suboptimal. It follows then that given V which satisfies (73), case 1 can be applied, which means that (72) must be satisfied.

Now note that if α varies with the value of z, it must be that given $z' \neq z''$,

$$\begin{aligned} J'\left(\sum_{n=1}^{N}\pi\left(\theta^{n}\right)\left(v\left(x\left(\theta^{n},z'\right)\right)+\beta V'\left(\theta^{n},z'\right)\right)\right) &= J'\left(\sum_{n=1}^{N}\pi\left(\theta^{n}\right)\left(v\left(x\left(\theta^{n},z''\right)\right)+\beta V'\left(\theta^{n},z''\right)\right)\right) \\ &= J'\left(V\right). \end{aligned} \tag{75}$$

This is because (15) and the weak concavity of $J(V)$ imply that

$$J\left(V\right) \geq \int_{0}^{1} J\left(\sum_{n=1}^{N}\pi\left(\theta^{n}\right)\left(v\left(x\left(\theta^{n},z\right)\right)+\beta V'\left(\theta^{n},z\right)\right)\right)dz, \tag{76}$$

and analogous arguments as used previous imply that conditional on z, social welfare must equal (74) which means that (76) must hold as an equality. However, if the value of

$$\sum_{n=1}^{N}\pi\left(\theta^{n}\right)\left(v\left(x\left(\theta^{n},z\right)\right)+\beta V'\left(\theta^{n},z\right)\right)$$

varies with z, for (76) to hold as an equality, it must be that (75) also holds. Combining (72) with (75), it follows that $J'\left(V'\left(\theta^{1},z\right)\right) > J'\left(V\right)\ \forall z$. ■

Lemma 13 *If $V=\overline{V}$, then the solution to* (14) − (21) *admits $V'\left(\theta^{1},z\right) < V\ \forall z$.*

Proof. Suppose that $V=\overline{V}$. Lemma 5 implies that the solution admits $P(z)=1\ \forall z$; part (i) of Lemma 8 implies that (20) binds $\forall z$; and the arguments in the proof of part (i) of Lemma 8 imply that the policymaker achieves a continuation value of $\overline{V}\ \forall z$.

Suppose it were the case that for some z, $V'\left(\theta^{1},z\right) \geq V$, which given part (i) of Lemma 6 as well as (21) implies that $V'\left(\theta^{n},z\right)=\overline{V}\ \forall n$. This implies that $J\left(V'\left(\theta^{n},z\right)\right)=u\left(\omega\right)/\left(1-\beta\right)$ by part (i) of Lemma 8. This also implies from (18) that $x\left(\theta^{n},z\right)=x\left(\theta^{n+1},z\right)+\theta^{n+1}-\theta^{n}$ $\forall n<N$. Therefore, since (20) binds, this means that means that $c\left(\theta^{n},z\right)=\omega\ \forall n$. Together with the fact that $\omega-i\left(z\right)+f\left(i\left(z\right)\right) \leq \omega-i^{*}+f\left(i^{*}\right)$ and that $c\left(\theta^{n},z\right)=\omega\ \forall n$, this means that $x\left(\theta^{n},z\right) \leq f\left(i^{*}\right)-i^{*}+\theta^{n}$. We now show that this previous relationship must hold with equality $\forall n$. Suppose it were the case that $x\left(\theta^{1},z\right) < f\left(i^{*}\right)-i^{*}+\theta^{1}$. Then this would imply from part (i) of Lemma 6 that $x(\theta^n, z) < f(i^*) - i^* + \theta^n$ so that

$$\overline{V} = \sum_{n=1}^{N} \pi(\theta^n) v(x(\theta^n, z)) / (1-\beta) < \sum_{n=1}^{N} \pi(\theta^n) v(f(i^*) - i^* + \theta^n) / (1-\beta). \tag{77}$$

However, given the arguments in the proof of Lemma 1, there exists a sustainable equilibrium which provides a welfare equal to the right hand side of (77). This however contradicts the fact that $\overline{V}$ corresponds to the highest sustainable welfare for the policymaker. Therefore, it is necessary that $x(\theta^n, z) = f(i^*) - i^* + \theta^n \ \forall n$. Note that since conditional on z, the policymaker receives $\overline{V}$ whereas households receive $u(\omega)/(1-\beta)$, a solution for which the elements of α do not depend on the realization of z exists.

We can focus on such a solution and we can show that this solution is suboptimal because it is possible to make households strictly better off while leaving the policymaker as well off. In order to establish this, we first establish the following upper bound on $J'(\overline{V})$ which must hold given that $x(\theta^n, z) = f(i^*) - i^* + \theta^n$ and $V'(\theta^n, z) = \overline{V}$ for all n and all z at $V = \overline{V}$. Define $J'(\overline{V}) = \lim_{\epsilon \to 0^+} (J(\overline{V}) - J(\overline{V} - \epsilon))/\epsilon$. Then it must be that

$$J'(\overline{V}) \leq -u'(\omega) / \left(\sum_{n=1}^{N} \pi(\theta^n) v'(f(i^*) - i^* + \theta^n) \right). \tag{78}$$

To see why this is the case, consider the following potential solution starting from $V = \overline{V} - \epsilon$ for $\epsilon > 0$ arbitrarily small. Let $P(z) = 1 \ \forall z$, $V'(\theta^n, z) = \overline{V} \ \forall n, z$, and $i(z) = i^* \ \forall z$. Moreover, $\forall n, z$, let $x(\theta^n, z) = f(i^*) - i^* + \theta^n - \varepsilon(\epsilon)$ for $\varepsilon(\epsilon)$ which satisfies

$$\epsilon = \sum_{n=1}^{N} \pi(\theta^n) (v(f(i^*) - i^* + \theta^n) - v(f(i^*) - i^* + \theta^n - \varepsilon(\epsilon))). \tag{79}$$

It is straightforward to verify that the conjectured solution satisfies all of the constraints of the problem for sufficiently small ϵ. Moreover, since such a solution is always feasible, it implies that

$$\frac{J(\overline{V} - \epsilon) - J(\overline{V})}{\epsilon} \geq \frac{u(\omega + \varepsilon(\epsilon)) - u(\omega)}{\epsilon}.$$

Taking the limit of both sides of the above inequality as ϵ approaches 0 implies (78).

Given the bound in (78), we can now show that the proposed solution at $V = \overline{V}$ is suboptimal. To see why, consider the following perturbation. Suppose that $x(\theta^n, z)$ were increased by $dx^n = \epsilon > 0$ arbitrarily small for all $n < N$. Moreover, suppose that $V'(\theta^n, z)$ were reduced by some $dV^n(\epsilon)$ which satisfies

$$dV^n(\epsilon) = \frac{1}{\beta} \sum_{n=1}^{N} \pi(\theta^n) (v(f(i^*) - i^* + \theta^n + \epsilon) - v(f(i^*) - i^* + \theta^n)) \tag{80}$$

for all $n < N$. Finally, suppose that $x\left(\theta^N, z\right)$ were decreased by some $dx^N(\epsilon)$ which satisfies:

$$\begin{aligned} v(f(i^*) - i^* + \theta^N - dx^N(\epsilon)) &= v(f(i^*) - i^* + \theta^N + \epsilon) \\ &\quad - \sum_{n=1}^{N} \pi(\theta^n)\left(v\left(f(i^*) - i^* + \theta^n + \epsilon\right) - v\left(f(i^*) - i^* + \theta^n\right)\right). \end{aligned}$$

Note that from the above $dx^N(\epsilon) > 0$. It can be verified that the proposed perturbation continues to satisfy all of the constraints of the problem. In order that this perturbation not strictly increase the welfare of households as ϵ approaches 0, it must be that:

$$-\sum_{n=1}^{N-1} \pi(\theta^n) u'(\omega) + \pi\left(\theta^N\right) u'(\omega) \lim_{\epsilon>0, \epsilon\to 0} \frac{dx^N(\epsilon)}{\epsilon} - \beta \sum_{n=1}^{N-1} \pi(\theta^n) J'\left(\overline{V}\right) \lim_{\epsilon>0,\epsilon\to 0} \frac{dV^n(\epsilon)}{\epsilon} \leq 0$$

Substituting (78) and (80) into the above implies that $\pi(\theta^N)u'(\omega)\lim_{\epsilon>0,\epsilon\to 0} \frac{dx^N(\epsilon)}{\epsilon} \leq 0$, which is a contradiction since $dx^N(\epsilon)$ is strictly positive. ■

We now go back to the proof of part (ii) and use the two previous lemma to show that $\underline{V}'$ does not exist so as to create a contradiction.

Step 1. Note that if $V \in \Psi$ so that the solution to (14) given V admits $i(z) = i^*$, it must be the case that $P(z) = 1\ \forall z$, that $V > V_0$, and that (19) is slack $\forall z$ so that Lemmas 12 and 13 can apply. To see why, suppose first it were the case that $P(z) = 0$ for some positive measure z. Then Lemma 5 would imply that $V < V_0$, leading to a contradiction of the fact that $\underline{V}' \geq V_0$. Furthermore, suppose it were the case that (19) is not slack so that the Lagrange multiplier $\psi(\theta^n, z) > 0$ for some some n. Then one can apply the arguments of step 2 in the proof of part (i) to show that this would imply that $i(z) < i^*$. Therefore, $\psi(\theta^n, z) = 0\ \forall n$ so that (19) is slack.

Step 2. Starting from any given $V_{t+k} \in \Psi$, $\Pr\{V_{t+k+1} < V_{t+k}\} \geq \pi\left(\theta^1\right)$. If $V_{t+k} = \overline{V}$, then this follows directly from Lemma 13 since $V_{t+k+1} < V_{t+k}$ if $\theta_{t+k} = \theta^1$. If instead $V_{t+k} \in \left(V_0, \overline{V}\right)$, it then follows from the arguments in Lemma 12. Therefore, in this case, $V_{t+k+1} < V_{t+k}$ if $\theta_{t+k} = \theta^1$.

Step 3. We show in this step that for any $V_{t+k} \in \Psi$, there exists a finite l such that if state 1 is repeated l times consecutively, then $V_{t+k+l} < \underline{V}'$. This step thus contradicts the fact that $\underline{V}'$ is the infinum of long run continuation values in Ψ, which completes the proof. To see why this is true, let $g(V)$ correspond to the highest realization of $V'\left(\theta^1, z\right) \in \Psi$ in the solution to the problem with state V. $g(V)$ is a continuous correspondence given the continuity of the objective function and the compactness and continuity of the constraint set. It follows that in a sequence under which θ^1 is repeated l times, $V_{t+k+l} \leq g\left(V_{t+k+l-1}\right) < V_{t+k+l-1}$, where we have used step 2. Suppose by contradiction that there is no finite l such that $V_{t+k+l} < \underline{V}'$. This means that $\lim_{l\to\infty} V_{t+k+l} = V' \geq \underline{V}'$ for some $V' \in \Psi$. This implies that $\lim_{l\to\infty} g\left(V_{t+k+l-1}\right) \geq V' \geq \underline{V}'$. However, given the continuity of $g(\cdot)$, this implies that

$g(V') \geq V'$, which contradicts the fact that $g(V) < V$ for all $V \in \Psi$.■

D.2 Proof of Proposition 3

Suppose that (27) holds and suppose that $\lim_{t\to\infty} \Pr\{P_{t+k} = 1\ \forall k\} > 0$, so that a permanent dictator emerges following some positive measure of histories. Let $\{V_t\}_{t=0}^{\infty}$ correspond to the stochastic sequence of continuation values under some such history, which must includes within it at least one convergent stochastic subsequence, where this follows from the fact that V_t is bounded. With some abuse of notation, let Ψ correspond to the set of all limiting values of convergent stochastic subsequences of $\{V_t\}_{t=0}^{\infty}$. It follows that $V \in \Psi$, that the solution to (14) $-$ (21) admits $V'(\theta, z) \in \Psi\ \forall\theta$ and $\forall z$ with positive measure since $V'(\theta, z)$ must itself be within the set of limits of stochastic subsequences. In order to proceed, we establish that for any $V \in \Psi$, $i(z) = i^*\ \forall z$. After this is established, the arguments in part (ii) of the proof of Proposition 2 can be used to show that this is not possible, completing the proof.

Given V, (18) for $\widehat{\theta} = \theta^1$ together with the fact that $x(\theta^1, z) \geq 0$ implies that conditional on z and θ,

$$v(x(\theta^n, z)) + \beta V'(\theta^n, z) \geq v(\theta^n - \theta^1) + \beta V'(\theta^n, z). \tag{81}$$

Since replacement never takes place in the future, forward iteration on (81) implies that

$$v(x(\theta^n, z)) + \beta V'(\theta^n, z) \geq v(\theta^n - \theta^1) + \beta \frac{\sum_{l=1}^{N} \pi(\theta^l)\, v(\theta^l - \theta^1)}{1-\beta}. \tag{82}$$

Together with (27), (82) implies that (19) holds as a strict inequality for all θ^n. To see why, from part (iv) of Lemma 6, it is sufficient to verify that (19) is slack if $\theta = \theta^1$. Given that that $i(z) \leq i^*$ from Lemma 11, it is clear that this is the case given (27) and (82). Since (19) holds as a strict inequality for all θ^n, the Lagrange multiplier $\psi(\theta^n, z)$ for constraint (19) must equal zero, which implies from (58) that $i(z) = i^*\ \forall z$. This establishes that if a permanent dictator emerges following some positive measure of histories, then the stochastic sequence $\{i_t\}_{t=0}^{\infty}$ converges to i^* for a positive measure of histories. However, this contradicts part (ii) of Proposition 2.■

D.3 Proof of Proposition 4

In order to prove this result, we establish the following preliminary lemmas.

Lemma 14 *$\exists V^* \in [V_0, \overline{V})$ s.t. if $V \geq V^*$, the solution to (14) $-$ (21) admits $i(z) = i^*\ \forall z$.*

Proof. We first establish that if $V = \overline{V}$, then (19) is a strict inequality and $i(z) = i^*$. Consider the solution given $V = \overline{V}$. We can use the same arguments as in the proof of Lemma 7 so as to perturb the solution so as to let $C_{n+1,n} = 0$, preserving the level of investment. The same arguments as those used in the proof of Lemma 13 imply that there is no replacement and policymakers receives a continuation value of $\overline{V}$ conditional on the realization of z. Using (18), it follows that for any $n^* \in \{1, ..., N\}$,

$$\begin{aligned}\overline{V} &= \sum_{n=1}^{N} \pi\left(\theta^n\right)\left(v\left(x\left(\theta^n, z\right)\right) + \beta V'\left(\theta^n, z\right)\right) \\ &\leq \sum_{n=1}^{n^*} \pi\left(\theta^n\right)\left(v\left(x\left(\theta^{n^*}, z\right)\right) + \beta V'\left(\theta^{n^*}, z\right)\right) + \sum_{n=n^*+1}^{N} \pi\left(\theta^n\right)\left(v\left(x\left(\theta^n, z\right)\right) + \beta V'\left(\theta^n, z\right)\right).\end{aligned}$$

Using part (i) of Lemma 6 together with the fact that $V'\left(\theta^{n^*}, z\right) \leq \overline{V}$, it follows that the above condition reduces to

$$\overline{V} \leq \frac{\sum_{n=1}^{n^*} \pi\left(\theta^n\right) v\left(x\left(\theta^{n^*}, z\right)\right) + \sum_{n=n^*+1}^{N} \pi\left(\theta^n\right) v\left(x\left(\theta^{n^*}, z\right) + \theta^n - \theta^{n^*}\right)}{1-\beta}. \tag{83}$$

Given Lemma 1, it is necessary that

$$\overline{V} \geq \frac{\sum_{n=1}^{N} \pi\left(\theta^n\right) v\left(f\left(i^*\right) - i^* + \theta^n\right)}{1-\beta}, \tag{84}$$

since a sustainable equilibrium providing the right hand side of (84) to the incumbent exists. (83) and (84) thus imply that

$$x(\theta^n, z) \geq f\left(i^*\right) - i^* + \theta^1 \forall n. \tag{85}$$

Now consider the value of $\underline{V}$. Suppose that (19) holds as an equality, where by part (iv) of Lemma 6, this must be the case for $\theta^n = \theta^1$. Taking into account that (19) holds as an equality together with the fact that local downward informational constraints bind, it is the case that $\overline{V}$ can be rewritten as

$$\overline{V} = v(f\left(i\left(z\right)\right) + \theta^1) + \beta\underline{V} + \sum_{n=2}^{N}\left(\sum_{l=2}^{N}\pi(\theta^l)\right)\left(v(x(\theta^{n-1}, z) + \theta^n - \theta^{n-1}) - v(x(\theta^{n-1}, z))\right). \tag{86}$$

(84) and (85) together with the fact that $v\left(\cdot\right)$ is strictly concave and $i\left(z\right) \leq i^*$ from Lemma 11 imply that (86) simplifies to

$$\frac{\sum_{n=1}^{N} \pi\left(\theta^n\right) v\left(f\left(i^*\right) - i^* + \theta^n\right)}{1-\beta} \leq v(f\left(i^*\right) + \theta^1\right) + \beta\underline{V} + \Gamma$$

for Γ defined in (29). However, if $N = 2$, it can be shown by some algebra that this violates Assumption 2 and if $N > 2$, this violates (28).

Therefore, if $V = \overline{V}$, then the solution admits (19) as a strict inequality and therefore $i\left(z\right) = i^*$. By standard arguments, the solution is continuous in V, which implies that if (19) is a strict inequality in the solution to (14) $-$ (21) given $V = \overline{V}$, then there exists some $V^* < \overline{V}$

such that (19) is a strict inequality in the solution to (14) − (21) given $V > V^*$ and therefore $i(z) = i^*$. ■

Lemma 15 *Consider the solution to* (14) − (21) *given* $V \in [V_0, \overline{V})$*. Then it must be that the solution admits either* $V'(\theta^N, z) = \overline{V}$ *or* $J'(V'(\theta^N, z)) < J'(V)\ \forall z$*.*

Proof. We pursue a similar strategy to that used in the proof of Lemma 12. Consider first the case in which the solution α does not vary with the realized value of z. Suppose that $V'(\theta^N, z) < \overline{V}$ and suppose by contradiction that $J'(V'(\theta^N, z)) \geq J'(V)\ \forall z$. Part (i) of Lemma 6 together with the weak concavity of $J(V)$ then implies that $J'(V'(\theta^n, z)) \geq J'(V)\ \forall n$ and $\forall z$, where we have used the fact that α does not depend on z. From Lemma 7, one can perturb such a solution without changing households' welfare and continuing to satisfy the constraints of the problem by changing the values of $V'(\theta^n, z)$ so that $C_{n+1,n} = 0$ for all $n < N$. Note that this perturbation weakly decreases the value of $V'(\theta^N, z)$ so that it remains the case that $J'(V'(\theta^N, z)) \geq J'(V)\ \forall z$.

Now consider such a solution. (59) together with (60) then imply that

$$\sum_{n=1}^{N} \pi(\theta^n) J'(V'(\theta^n, z)) = J'(V) - \pi(\theta^1)\psi(\theta^1, z) - \pi(\theta^1)\underline{\mu}(\theta^1, z) \tag{87}$$

where we have used the fact that $\eta(z) = 0$ since otherwise (68) binds $\forall z$ so that $J(V) = u(\omega) / (1 - \beta)$, and where we have used Lemma 6 together with the fact that $J'(V'(\theta^n, z)) \geq J'(V)\ \forall n$. It thus follows from the above condition that $\psi(\theta^1, z) = 0$ so that (19) can be ignored. Analogous arguments as in the proof of Lemma 12 can then apply and be used to argue that in this case that $J'(V'(\theta^N, z)) < J'(V)\ \forall z$, leading to a contradiction. Moreover, analogous arguments as those used in the proof of Lemma 12 for the case in which α varies with z also apply. ■

Proof of part (i). We begin by looking at investment dynamics and proceed in an analogous fashion as in the proof of Proposition 2. Suppose it were the case that $\lim_{t \to \infty} \Pr\left\{i_{t+k} = \widehat{i}\ \forall k\right\} > 0$ for some $\widehat{i}$. By Proposition 2, it is necessarily the case that $\widehat{i} \neq i^*$. This would imply that there exist a stochastic sequence $\{i_t\}_{t=0}^{\infty}$ with positive measure which converges to $\widehat{i}$. Associated with such a stochastic sequence is a stochastic sequence $\{V_t\}_{t=0}^{\infty}$ which must includes within it at least one convergent stochastic subsequence, where this follows from the fact that V_t is bounded. Let Ψ correspond to the entire set of all limiting values of convergent stochastic subsequences of $\{V_t\}_{t=0}^{\infty}$ for every single stochastic sequence $\{i_t\}_{t=0}^{\infty}$ which converges to $\widehat{i}$. It follows by the continuity of the policy function that the solution to (14) − (21) starting from some $V \in \Psi$ yields a solution which admits $i(z) = \widehat{i}\ \forall z$. Moreover, it must be that $V'(\theta, z) \in \Psi\ \forall \theta$ and $\forall z$ with positive measure since $V'(\theta, z)$ must itself be within the set of limits of stochastic subsequences. Let $\overline{V}'$ correspond to the suprenum of continuation values in Ψ. It is clear that $\overline{V}' < V^*$ for V^* defined in Lemma 14, since if this were not the case, this would imply that $i(z) = i^* \neq \widehat{i}$ for some $V \in \Psi$, leading to a contradiction. The strategy of our proof is to show that $\overline{V}'$ does not exist so as to create a contradiction. We proceed in the following steps.

Step 1. Starting from any given $V_{t+k} \in \Psi$, $\Pr\{V_{t+k+1} > V_{t+k}\} \geq \pi\left(\theta^N\right)$. If $V_{t+k} \in [V_0, \overline{V})$, then this follow directly from Lemma 15. If instead $V_{t+k} < V_0$, then independently of the realization of z, the policymaker choosing policies–whether it is the previous incumbent or the new replacement policymaker–receives an expected continuation value equal to V_0. As such, the arguments of Lemma 15 starting from $V = V_0$ establish that $V_{t+k+1} > V_0 > V_{t+k}$ if $\theta_{t+k} = \theta^N$.

Step 2. We show in this step that for any $V_{t+k} \in \Psi$, there exists a finite l such that if state 1 is repeated l times consecutively, then $V_{t+k+l} > \overline{V}'$. This step thus contradicts the fact that $\overline{V}'$ is the suprenum of long run continuation values in Ψ, which completes the proof. To see why this is true, with some abuse of notation, let $g(V)$ correspond to the lowest realization of $V'\left(\theta^N, z\right) \in \Psi$ in the solution to the problem with state V. $g(V)$ is a continuous correspondence given the continuity of the objective function and the compactness and continuity of the constraint set. It follows that in a sequence under which θ^N is repeated l times, $V_{t+k+l} \geq g\left(V_{t+k+l-1}\right) > V_{t+k+l-1}$, where we have used step 1. Suppose by contradiction that there is no finite l such that $V_{t+k+l} > \overline{V}'$. This means that $\lim_{l\to\infty} V_{t+k+l} = V' \leq \overline{V}'$ for some $V' \in \Psi$. This implies that $\lim_{l\to\infty} g\left(V_{t+k+l-1}\right) \leq V' \leq \overline{V}'$. However, given the continuity of $g(\cdot)$, this implies that $g(V') \leq V'$, which contradicts the fact that $g(V) > V$ for all $V \in \Psi$. This completes the proof that $\lim_{t\to\infty} \Pr\left\{i_{t+k} = \widehat{i}\ \forall k\right\} = 0$.

Proof of part (ii). Proving the second part of the proposition is straightforward given the dynamics of continuation values. Suppose it were the case that $\lim_{t\to\infty} \Pr\{x_{t+k}(\theta) = \widehat{x}(\theta)\ \ \forall k, \theta\} = 1$ for some $\widehat{x}(\theta)$. Using analogous reasoning as in the proof of part (i), one can define, with some abuse of notation, the set Ψ of limiting continuation values. It necessarily follows that the solution to (14) $-$ (21) starting from some $V \in \Psi$ yields a solution which admits $x(\theta, z) = \widehat{x}(\theta)\ \forall\theta$ and $\forall z$. Moreover, it must be that $V'(\theta, z) \in \Psi\ \forall\theta$ and $\forall z$ with positive measure since $V'(\theta, z)$ must itself be within the set of limits of stochastic subsequences. This leaves two cases to consider.

Case 1. Suppose that $\lim_{t\to\infty} \Pr\{P_{t+k} = 1\ \forall k\} = 1$ so that there is no replacement in the long run. Given (15), this requires $\lim_{t\to\infty} \Pr\left\{V_{t+k} = \widehat{V}\ \forall k\right\} = 1$ for some $\widehat{V} > V_0$ so that Ψ is a singleton. However, Lemmas 12 and 13 together imply that $\Pr\{V_{t+k+1} < V_{t+k}\} \geq \pi\left(\theta^1\right)$ so that $\{V_t\}_{t=0}^{\infty}$ cannot converge to a single point since it must necessarily decrease.

Case 2. Suppose that $\lim_{t\to\infty} \Pr\{P_{t+k} = 1\ \forall k\} < 1$ so that there is replacement in the long run. We can show that in this case, it must be that $\sum_{n=1}^{N} \pi(\theta^n)\left(v\left(x\left(\theta^n, z\right)\right) + \beta V'\left(\theta^n, z\right)\right) = V_0$ starting from all $V_{t+k} \in \Psi$. To see why, suppose this were not the case and consider the solution (14) $-$ (21) starting from some $V \in \Psi$. Use Lemma 7 so as to perturb such a solution without changing households' welfare and continuing to satisfy the constraints of the problem by changing the values of $V'(\theta^n, z)$ so that $C_{n+1,n} = 0$ for all $n < N$. Given that

$x(\theta,z) = \widehat{x}(\theta)$ $\forall\theta$ and $\forall z$, it follows that satisfaction of (15) and (18) uniquely determines $\{V'(\theta^n,z)\}_{n=1}^N$ $\forall z$, where $V'(\theta^n,z)$ is strictly increasing in V. From (19) and (58), it follows that $i(z)$ is weakly larger the larger the value of $V'(\theta^1,z)$ which means that $i(z)$ is weakly increasing in V. Lemma 11 implies that $i(z) \leq i^*$, which combined with (16), furthermore implies that the value of $\sum_{n=1}^N \pi(\theta^n)\, u(c(\theta^n,z))$ is weakly increasing in V. However, combined with the fact that $V'(\theta^n,z)$ is strictly increasing in V, this means that for $J(V)$ to be non-increasing in the range of $V \in \Psi$–which is established in Lemma 4–it must be that the set Ψ is either a singleton, which is ruled out in case 1, or that $\Psi \subset [\underline{V}, V_0]$, in which case $\sum_{n=1}^N \pi(\theta^n)\,(v(x(\theta^n,z)) + \beta V'(\theta^n,z)) = V_0$. However, the arguments in step 1 in the proof of part (i) imply that starting from $V_{t+k} = V_0$, $\Pr\{V_{t+k+1} > V_{t+k}\} \geq \pi(\theta^N)$, which means that it is not possible for $\Psi \subset [\underline{V}, V_0]$.■