Political Distortions and Endogenous Turnover*

Laurence Ales† Pricila Maziero‡ Pierre Yared§

May 4, 2012

Abstract

We study the best sustainable equilibrium in an economy with non-benevolent policymakers who lack commitment and have private information. In this environment, households discipline policymakers by threatening to remove them from power. Policymakers are never replaced and there are no distortions to production if there is perfect information. We present three results which emerge once private information is introduced. First, we generalize the endogenous turnover result of Ferejohn (1986): an incumbent always faces a positive probability of replacement starting from the initial date in power since this provides him incentives to not privately rent-seek. Second, we show that the presence of endogenous turnover generates distortions to production. These distortions emerge so as to limit the resources which can be expropriated by a policymaker facing replacement. Finally, policymakers are always replaced and distortions to production never vanish in the long run.

Keywords: Optimal Taxation, Political Economy, Asymmetric and Private Information
JEL Classification: H21, P16, D82

*We would like to thank Aleh Tsyvinski and Mike Golosov and participants at the SED conference for comments.

†Tepper School of Business, Carnegie Mellon University. e-mail: ales@cmu.edu.
‡The Wharton School, University of Pennsylvania. email: maziero@wharton.upenn.edu.
§Columbia University, Graduate School of Business. e-mail: pyared@columbia.edu.

1 Introduction

In practice, government policies are not chosen by benevolent planners, but by policymakers who are partly motivated by political rents and by the desire to preserve power.[1] Policymakers generally cannot commit to the policies which they promise their citizens, since, once in office, they can always choose to pursue their self-interest. Moreover, the rents which they extract in office may be private information which cannot be perfectly observed by citizens. As an example, policymakers with little scrutiny can often overpay for the procurement of a public good as a means of accessing additional funds for themselves or their supporters.[2]

In this paper we take these considerations into account to develop a dynamic political economy model which combines two frameworks. Our starting point is the classic Ferejohn (1986) model of political accountability under asymmetric information. We combine this framework with a model of a dynamic production economy with rent-seeking along the lines of some recent work on the political economy of dynamic fiscal policy which assumes perfect information.[3] Our analysis therefore generalizes the environment of Ferejohn (1986) to a setting in which there is investment and production and in which policymakers and citizens pursue fully history dependent (non-Markovian) strategies.[4] Our setting features three key frictions which motivate our analysis: Policymakers are non-benevolent, they lack commitment, and they have private information. The purpose of our analysis is to understand the effect of these three frictions for the dynamics of production and policy.

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. There is aggregate uncertainty in the form of an additive shock to the government's budget constraint, where this captures a shock to the cost of public

[1]For a discussion of the self-interested behavior of politicians and the implications for corruption and public goods provision, see Acemoglu (2003), Acemoglu, Johnson, and Robinson (2005), Acemoglu and Verdier (2000), Banerjee (1997), Buchanan and Tullock (1967), North et al. (1981), Persson and Tabellini (2000), and Shleifer and Vishny (1993).

[2]As an example, Olken (2007) documents the over-reporting of project costs by government officials using a field experiment in Indonesia. As another example, there is evidence that there was significant fraud and that the U.S. government over-paid for relief in the response to Hurricane Katrina (see GAO-Report (2007), Eaton (2006), and Lipton (2005)).

Alternatively, policymakers may sometimes be able to secretly divert government royalties for themselves. Caselli and Michaels (2009), for example, report that large oil output tends to be associated with an increase in instances of alleged illegal activities by mayors in Brazilian municipalities. There is some evidence government royalties from oil and gas may even be misreported in the U.S. (see GAO-Report (2010)).

[3]Studies which consider the policy implications of rent-seeking but ignore issues of private information include–though by no means are limited to–Acemoglu, Golosov, and Tsyvinski (2008, 2010a,b), Aguiar, Amador, and Gopinath (2009), Azzimonti (2010), Battaglini and Coate (2008), Caballero and Yared (2010), Krusell and Rios-Rull (1999), Song, Storesletten, and Zilibotti (2009), and Yared (2010).

[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. Barro (1973) shows that the threat of removal need only be used off the equilibrium path if there is full information. See also Banks and Sundaram (1998), Besley (2006), Egorov (2009), Fearon (2010), and Persson and Tabellini (2000) for extensions.

spending or to the value of government royalties. 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 constraints of both limited commitment and private information on the side of the policymaker. Our benchmark environment focuses on the role of private information and considers a setting in which the best allocation for households emerges under full information. More specifically, absent private information, the level of investment is efficient, and the threat of off-equilibrium replacement is sufficient to induce a policymaker to extract zero rents. Therefore, in the efficient sustainable equilibrium under full information, there is no turnover and there are no distortions to production.

We present three results which emerge once private information is introduced. Our first result generalizes the endogenous turnover result of Ferejohn (1986). We find that an incumbent always faces a positive probability of *equilibrium* replacement starting from the initial date in power. The reasoning behind this result is that the threat of replacement must be exercised along the equilibrium path in order to provide the policymaker with the right incentives to truthfully reveal the state of the economy to citizens. More specifically, the optimal contract between citizens and the policymaker rewards the policymaker in the future whenever the aggregate shock slackens the government budget constraint and punishes the policymaker in the future whenever the aggregate shock tightens the government budget constraint. This induces the policymaker to not lie to citizens and privately misappropriate funds whenever the aggregate shock slackens the government budget constraint. Whereas transfering future rents to reward policymakers is costly for citizens, replacing policymakers in the future to punish them is costless for citizens. For this reason, from an ex-ante perspective, it is efficient for citizens to punish new incumbents with possible replacement in order to induce them to reveal their private information. This result generalizes the endogenous turnover result of Ferejohn (1986) to an economy in which the flow payoff of holding political power is not exogenous but endogenous to economic policy; where production is determined by optimizing households; and where policymakers and citizens choose fully history dependent strategies associated with the best sustainable equilibrium.

Our second main result regards the joint determination of endogenous turnover and economic distortions to production. More specifically, we find that production is always distorted whenever there is a positive probability of turnover. The reasoning is as follows. Endogenous turnover in equilibrium disciplines policymakers by inducing them to not privately divert funds to themselves. Nonetheless, there is a limit to the extent to which replacement is a useful tool, since policymakers expecting future replacement may lack the incentives to pursue prescribed policies, and may decide instead to fully expropriate households. As such, distortions to production arise endogenously since they limit the available resources which can be expropriated by the policymaker and therefore reduce the equilibrium rents which need to be transfered to the policymaker in order to prevent expropriation.

Our final result regards the characterization of the dynamics of rents, turnover, and investment. Specifically, we find that, in the long run, turnover persists so that incumbent policymakers are always replaced and distortions to production *never* vanish. The intuition for our result is that, even though a policymaker can be temporarily rewarded with an increase in rents and an increase in tenure following a sequence of shocks which slacken the government budget constraint, eventually, all policymakers will be subject to a sequence of shocks which tighten the government budget. Optimal incentive provision requires the welfare of these policymakers to decline following these shocks, and this reduces the rents which they receive and ultimately leads to their turnover and to the emergence of distortions to production. Importantly, this result arises as a consequence of optimality and *not* feasibility since allocations in which the incumbent remains in power forever and truthfully reports the economic state to citizens are sustainable in our environment; however, they are suboptimal since they do not entail enough risk-sharing between households and the policymaker.[5]

Note that this final result is a consequence of the competing frictions in our framework. On the one hand, the constraint of limited commitment puts upward pressure on the policymaker's future welfare. This emerges because *backloading* is optimal: a policymaker who does not deviate in the present is rewarded in the future with higher welfare, and hence faces longer tenure, higher rents, and smaller economic distortions.[6] On the other hand, the constraint of private information on the side of the policymaker puts both upward *and* downward pressure on the policymaker's welfare. This serves as the most efficient means of providing him with incentives to reveal the state of the economy to the citizens (so that he does not privately rent-seek) while simultaneously smoothing his consumption and that of the citizens as much as possible. While these forces compete in a dynamic framework, there is a natural upper limit to the welfare which the policymaker can achieve without disincentivizing households from investing altogether, and this limits the extent to which the policymaker can be rewarded. As such, the downward pressure ultimately dominates, so that the policymaker's rents must decline with positive probability until he is punished with replacement.[7] Note that this result is in sharp contrast to the related model of Acemoglu, Golosov, and Tsyvinski (2008, 2010a,b). They find that if the policymaker and the citizens have the same discount factor (as

[5]Note that we do not make assumptions regarding the size of this uncertainty. Our results hold for any arbitrarily small value of uncertainty.

[6]This insight emerges in the work of Acemoglu, Golosov, and Tsyvinski (2008, 2010a,b), Kocherlakota (1996), Ray (2002), and Thomas and Worrall (1988, 1990), among others.

[7]This insight also emerges in economies with risk sharing under private information (e.g., Thomas and Worrall (1990) and Atkeson and Lucas (1992). More specifically, these models predict that immiseration takes place as a consequence of optimal incentive provision. Our result is related to the insights from this work, though the reasoning is related to the presence of double-sided risk-sharing and not to the Inada condition on preferences. See Phelan (1998) for a discussion of the role of the Inada condition in this framework.

in our framework), then the policymaker has permanent tenure and distortions to production vanish in the long run. The key difference between our work and this work is that we allow for asymmetric information between the policymaker and the citizens, and this leads to the different characterization of the long run dynamics of the economy.[8]

Our paper is most closely related to the recent work in dynamic political economy which considers the effect of political frictions for production and policy.[9] We focus on the role of private information which is not present in this previous work. Whereas previous work features either exogenous political turnover or no political turnover, we show that private information generates endogenous political turnover which in turn creates economic distortions. Our paper is also related to several lines of research which consider the role of private government information (e.g., Athey, Atkeson, and Kehoe (2005), Amador, Werning, and Angeletos (2006), and Sleet (2001, 2004)). 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. In this respect, our paper is complementary to the work of Rogoff and Sibert (1988) and Rogoff (1990) who consider an economy in which office-driven policymakers have private information about their competency. 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. This allows us to characterize how society should optimally structure replacement rules in the best sustainable equilibrium to minimize rent-seeking.[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 summarizes our results. Section 5 provides a numerical simulation of the model and some additional results regarding transitional dynamics. 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

[8] In addition, note that in contrast to this work, distortions in our environment only emerge once private information is introduced since it generates a need for equilibrium turnover, which in turn tightens the constraint of limited commitment on the policymaker. See Section 4.4 for further discussion.

[9] See the work cited in footnote 3.

[10] In other words, this work considers the role of prospective voting, whereas our work considers the role of retrospective voting. Drazen and Ilzetzki (2011) also consider the role of private information in an environment with competing parties, though in contrast to our work they do not characterize optimal sustainable policies or allow for replacement.

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 $\theta^n > \theta^{n-1}$. 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. $u(\cdot)$ is strictly increasing and strictly concave in c_t with $\lim_{c \to 0} u_c(\cdot) = \infty$ and $\lim_{c \to \infty} u_c(\cdot) = 0$. Households enter every period with a fixed endowment $\omega > 0$. They decide how of 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:

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

where $\tau_t(y^t)$ is a function which 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 $\tau_t(y^t) \leq y_t$, meaning that the government cannot impose a tax on production which exceeds one hundred percent. Note 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:

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

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 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:

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

where we have taken into account that since households are identical, the government's aggregate tax revenue equals the individual tax burden $\tau_t(y^t)$. θ_t represents aggregate uncertainty which is determined after investment is undertaken and before policies $\tau_t(y^t)$ are chosen. It captures a shock to the cost of public spending or to the value of government royalties.

The resource constraint of the economy implied by (2) and (4) is:

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

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. We assume that $f(i^*) - i^* + \theta^1 > 0$, so that is feasible to pay the policymaker positive rents under all shocks while providing households with their outside option ω.

The most important feature of this setting is that while the entire society observes the policy $\tau_t(y^t)$, 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 $k \in J$). To reduce notation, we let $P_t = \{0, 1\}$ correspond to the to the decision of whether or not keep an incumbent at date t.

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 $u(c_t) = -\infty$.[11]

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 or protests. We simplify the discussion by assuming that the decision is taken by the same single representative citizen in every period.[12]

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,

[11] 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.

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

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 is 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 politician 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.

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 tax policies 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 ξ assigns a level of investment at every h_t^1. The incumbent policymaker's strategy ν 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 $\xi|_{h_t^1}$ and $\nu|_{h_t^2}$ analogously.[13]

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 $\{\xi, \nu\}$. A representative household's investment sequence ξ solves the representative household's problem if at every h_t^1, the continuation investment sequence $\xi|_{h_t^1}$ maximizes household welfare given $\{\Upsilon, \nu\}$ and given the household's budget constraint. The incumbent politician's

[13]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.

strategy ν solves the incumbent politician's problem if, at every h_t^2, the continuation strategy $\nu|_{h_t^2}$ maximizes the incumbent politician's welfare given $\{\Upsilon, \xi\}$ 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 $\{\Upsilon, \xi, \nu\}$ for which Υ solves the representative citizen's problem, ξ solves the household's problem, and ν solves the incumbent politician'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:

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

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

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

the welfare expected by the incumbent at the beginning of the stage game prior to the realization of the public signal z_t.[14] Moreover, define $J(q_t)$ analogously as the welfare of the households prior to the realization of z_t:

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

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

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

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

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

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

[14] Throughout the paper, we will refer to $V(q_{t+1})$ as $V(q_t, z_t, \theta_t)$.

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

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

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

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

Proof. See Appendix. ■

The intuition for Proposition 1 is as follows. The government has significant flexibility in choosing its non-linear tax instrument $\tau_t\left(y^t\right)$. 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 $u\left(\omega\right) / \left(1 - \beta\right)$. 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.[15]

Constraints (9) and (10) capture the incentive compatibility constraints on the side of the policymaker. They emerge from the fact that citizens can discipline a policymaker by punishing observable deviations by removing him from power. 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 $V((q_t, z_t, \widehat{\theta}))$ at $t+1$. Constraint (9) guarantees that he weakly prefers to choose the prescribed policy which provides him with rents equal to $x_t\left(q_t, z_t, \theta_t\right)$ at t and a continuation value of $V\left(\left(q_t, z_t, \theta_t\right)\right)$ 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.[16]

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

[15] Note that if taxes could not be history dependent and could only depend on y_t, then (11) would be replaced by $\sum_{\theta_t \in \Theta} \pi\left(\theta_t\right) u\left(c_t\left(q_t, z_t, \theta_t\right)\right) \geq u\left(\omega\right) / \left(1 - \beta\right)$. 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.

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

incumbent and any replacement policymaker.[17]

In order to guarantee the existence of a sustainable equilibrium, we make the following assumption for the remainder of our analysis.

Assumption 1 $\underline{V}$ *satisfies*

$$v\left(f\left(i^{*}\right)-i^{*}+\theta^{1}\right)+\beta\frac{\sum_{\theta\in\Theta}\pi\left(\theta\right)v\left(f\left(i^{*}\right)-i^{*}+\theta\right)}{1-\beta}>v\left(f\left(i^{*}\right)+\theta^{1}\right)+\beta\underline{V}. \quad (12)$$

Under Assumption 1, 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 such an equilibrium, 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 (7), (8), and (11). It also implies that $V(q_t) = \sum_{\theta\in\Theta}\pi(\theta)v(f(i^*) - i^* + \theta)/(1-\beta) > v(0)/(1-\beta)$ for all q_t and that $x_t(q_t, z_t, \theta_t) = x_t(q_t, z_t, \widehat{\theta}) + \theta_t - \widehat{\theta}$ for all (q_t, z_t, θ_t) and $\widehat{\theta}$. Therefore, (9) is satisfied. Moreover, by Assumption 1, (10) 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 1 implies that (10) is satisfied if $\theta_t = \theta^n$. Therefore, δ is supported by sustainable equilibrium strategies. ■

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

$$\max_{\delta\in\Lambda} E_0 \sum_{t=0}^{\infty}\beta^t u\left(c_t\left(q_t, z_t, \theta_t\right)\right), \quad (13)$$

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.

[17]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.

3.3 Recursive Representation of Best Sustainable Equilibrium

To facilitate the analysis, we provide a recursive formulation for (13). Define $\overline{J}$ as the utility attained under the solution to (13). Note that if the solution to (13) 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 (13) 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.[18]

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

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

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 to the incumbent policymaker. The recursive program is:

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

s.t.

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

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

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

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

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

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

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

[18] 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.

(14) takes into account that if $P(z) = 0$, the incumbent policymaker is replaced and households receive a continuation welfare $\overline{J}$.[19] 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 values $V'(\theta, z)$ is in the feasible range between $\underline{V}$ and $\overline{V}$.[20] The following lemma describes several important properties of $J(V)$.

Lemma 2 *$J(V)$ 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 $V \in [V_0, \overline{V}]$, (iii) and it is continuously differentiable in V for $V \in (\underline{V}, \overline{V})$.*

Proof. See Appendix. ■

Lemma 2 is useful since it facilitates our characterization of the best sustainable equilibrium and guides us in understanding the dynamics of the welfare of the policymaker. That $J(V)$ is 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. A subtle result embedded in this Lemma 2 is that $J'(V) = 0$ for $V \in (\underline{V}, V_0]$. The reason for this is that 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 z). Therefore, the welfare of households does not vary with V in this range.

4 Analysis

4.1 Full Information Benchmark

As a benchmark, we begin by considering the environment with full information, so that the households observe θ_t and x_t so that replacement decisions can be conditioned on the shock to the economy as well as the policies chosen by the policymaker. This corresponds to the

[19] This continuation welfare is associated with the replacement policymaker receiving a continuation value V_0.

[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.

solution to (14) which ignores (18). In this situation, all deviations by the policymaker from prescribed policies are observable and punished by replacement.

Throughout our analysis, we will refer to an economic distortion as a situation in which $f'(i) \neq 1$ so that the level of investment is not socially efficient. Note that if $f'(i) > 1$, then this will correspond to a situation in which $i < i^*$ so that there is under-investment relative to the socially efficient level. We consider a situation in which $\underline{V}$ satisfies the following assumption.

Assumption 2

$$\frac{v(0)}{1-\beta} \geq v(f(i^*) + \theta^N) + \beta \underline{V}$$

Assumption 2 implies that, under the efficient level of investment, a policymaker prefers to remain in power forever and to consume 0 rents versus fully expropriating households today and being thrown out today. We make this assumption because it has the following implication regarding the best sustainable equilibrium under full information.

Proposition 2 ***(full information)*** *Under full information under Assumption 2, the best sustainable equilibrium features no distortions ($i_t = i^*$), zero rents ($x_t = 0$), and no replacement ($P_t = 1$) for all t.*

Proof. See Appendix. ■

Proposition 2 states that the best allocation for households is achievable even under a self-interested policymaker with limited commitment. Such an allocation involves the minimal rents and the efficient level of investment. Even though he receives 0 rents, the policymaker does not expropriate the households because the cost of being thrown out of office if he does so is very large, and this is implied by Assumption 2.

The implication of Proposition 2 is that any inefficiencies which emerge in our setting must emerge because of the presence of imperfect information. Given that our focus is the effect of imperfect information, we preserve Assumption 2 for the remainder of our analysis in this section. In Section 4.4, we consider the extent to which our results can be generalized in environments in which $\underline{V}$ is large enough that Assumption 2 is violated.[21]

4.2 Equilibrium Dynamics under Imperfect Information

We consider an environment subject to imperfect information so that constraint (18) is taken into account. In this situation, the policymaker may be able to deviate from prescribed policies in an unobservable fashion. In order for this informational constraint to play a role, we assume that $\underline{V}$ is sufficiently high that it satisfies the following assumption.

[21] One can easily show that Assumption 2 implies Assumption 1. In Section 4.4, we relax Assumption 2 while preserving Assumption 1.

Assumption 3

$$\frac{v\left(0\right)-\sum_{n=2}^{N}\left(v\left(\theta^{n}-\theta^{n-1}\right)-v\left(0\right)\right)\left(1-\beta\sum_{k=n}^{N}\pi\left(\theta^{k}\right)\right)}{1-\beta}<v(f(i^{*})+\theta^{1})+\beta\underline{V}.\text{[22]} \tag{22}$$

Lemma 3 *If Assumption 3 does not hold, then the best sustainable equilibrium features no distortions ($i_t = i^*$), zero rents ($x_t = 0$), and* $\Pr\left\{P_{t+k}=1\ \forall k\geq 0\right\}=0$ *for all* t*.*

Proof. See Appendix. ■

Lemma 3 illustrates how, if $\underline{V}$ is sufficiently low so that Assumption 3 is violated, the constraint of imperfect information does not impose any welfare cost on households. This is because the value of remaining in power is sufficiently high for a policymaker that it is possible to induce him to follow prescribed policies without having to resort to paying him with rents. In equilibrium, whenever θ_t is high, households reward the policymaker in the future by letting him remain in power. Whenever θ_t is low households punish the policymaker in the future by removing him from power with positive probability. This illustrates the use of political replacement as a means of inducing good behavior on the side of the policymaker. For the remainder of our analysis, we impose Assumption 3. We now prove the first main result of the paper that states that a new incumbent always faces a positive probability of replacement.

Proposition 3 ***(turnover)*** *In the best sustainable equilibrium, all new incumbent policymakers face a positive probability of future replacement (i.e.,* $\exists\ k>0$ *such that* $\Pr\left\{P_{t+k}=0|P_t=0\right\}>0\ \forall t$*).*

Proof. See Appendix. ■

Together with Lemma 3, Proposition 3 generalizes the endogenous turnover result of Ferejohn (1986). As in Ferejohn (1986), an incumbent always faces a positive probability of *equilibrium* replacement starting from the initial date in power. This insight is generalized here to an economy in which the flow payoff of holding political payoff is not exogenous but endogenous to economic policy; where production is determined by optimizing households; and where policymakers and citizens choose fully history dependent strategies associated with the best sustainable equilibrium.

[22]Note that it can be verified that the set of values for $\underline{V}$ which satisfies Assumptions 2 and 3 is never empty if the following condition holds:

$$\left(v\left(f\left(i^{*}\right)+\theta^{N}\right)-v\left(f\left(i^{*}\right)+\theta^{1}\right)\right)\left(1-\beta\right)<\sum_{n=2}^{N}\left(v\left(\theta^{n}-\theta^{n-1}\right)-v\left(0\right)\right)\left(1-\beta\sum_{k=n}^{N}\pi\left(\theta^{k}\right)\right).$$

This is always true given the concavity of $v\left(\cdot\right)$ since

$$\begin{aligned}\sum_{n=2}^{N}\left(v\left(\theta^{n}-\theta^{n-1}\right)-v\left(0\right)\right) &\geq \sum_{n=2}^{N}\left(\frac{v\left(\theta^{N}-\theta^{1}\right)-v\left(0\right)}{\theta^{N}-\theta^{1}}\right)\left(\theta^{n}-\theta^{n-1}\right)\\ &> v\left(f\left(i^{*}\right)+\theta^{N}\right)-v\left(f\left(i^{*}\right)+\theta^{1}\right).\end{aligned}$$

The intuition for this result is as follows. As we discussed, the policymaker must be induced during the high shock to not privately deviate and appropriate the additional funds that would otherwise be transfered to the government under the low shock. This means that households provide the policymaker with higher welfare (a reward) from date $t+1$ onward following a high shock at t and a lower welfare (a punishment) from date $t+1$ onward following a low shock at t. Suppose by contradiction that a new incumbent were never replaced and that only rents were used to induce him to not privately deviate from equilibrium policy. In this situation, rents could be reduced along the equilibrium path without affecting incentives and this would be strictly beneficial for households. Specifically, it would be possible for households to reward the policymaker with fewer rents following a high shock and to punish the policymaker with a higher replacement probability following a low shock. Following replacement, a new incumbent would enter who would face the same low level of rents going forward. In other words, replacement is a cheaper tool for households to use in providing incentives to policymakers relative to rents. Therefore, turnover occurs along the equilibrium path.

A natural questions concerns the effect of endogenous turnover on economic distortions. The next proposition shows that in the best sustainable equilibrium, there are economic distortions whenever the probability of turnover is positive.

Proposition 4 ***(distortions)*** *In the best sustainable equilibrium, if there is turnover at t (i.e.,* $\Pr\{P_t = 0\} > 0$*), then there are distortions at t (i.e.,* $i_t < i^*$*).*

Proof. See Appendix. ■

The reasoning for this proposition is related to the intuition behind Proposition 3 that endogenous turnover must take place since replacement is a cheaper means of incentive provision relative to rent-seeking. The presence of limited commitment on the side of the policymaker constrains the use of replacement since an incumbent expected to be replaced in a future period may choose to publicly deviate from prescribed policies today and expropriate *all* of the resources in the economy. In this light, distortions to investment and to production facilitate incentive provision by limiting the level of resources which a policymaker expecting to be replaced in the future can acquire by deviating today. This makes possible to reduce the equilibrium rents paid to the policymaker today, and this is ex-ante efficient for citizens. Note that these distortions emerge not because taxation is assumed to be distortionary (e.g., lump sum taxes are allowed), but they arise endogenously in response to the presence of endogenous replacement. The insight that endogenous turnover can result in economic distortions does not arise in the work of Ferejohn (1986) since he does not consider an environment in which optimizing households invest inputs into production.

To see a heuristic proof of this argument, suppose it were the case that $i_t = i^*$ and suppose that there were two shocks θ^1 and $\theta^2 > \theta^1$. For simplicity, suppose further that the constraint of imperfect information (9) only binds under the high shock and is slack under the low shock. Suppose further that the constraint of limited commitment (10) is only relevant under the low shock. 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, suppose that households reduce investment by $\epsilon > 0$ arbitrarily small. This perturbation relaxes the right hand side of (10) by approximately $\epsilon v' \left(f\left(i^*\right) + \theta^1\right) f'\left(i^*\right)$. This perturbation allows for the reduction of rents to the policymaker under the low shock by approximately $\epsilon v' \left(f\left(i^*\right) + \theta^1\right) f'\left(i^*\right) / v' \left(x\left(\theta^1, z\right)\right)$, where $x\left(\theta^1, z\right)$ are the equilibrium rents he is receiving under the low shock. This perturbation continues to satify (9) under the high shock so that the incentive compatibility constraints are satisfied. Furthermore, household consumption under the low shock changes by approximately

$$-\left(f'\left(i^*\right) - 1\right)\epsilon + \epsilon v'\left(f\left(i^*\right) + \theta^1\right) f'\left(i^*\right) / v'\left(x\left(\theta^1, z\right)\right)$$

which exceeds 0 since $f'\left(i^*\right) = 1$. Therefore, distortions can make households strictly better off. Note that this insight emerges because of Assumption 3 which guarantees that the cost of being thrown out power is sufficiently low to the policymaker that it is not possible to induce him to not privately rent-seek without paying him rents along the equilibrium path. Therefore, one interpretation of Proposition 4 is that it implies that if rents are paid along the equilibrium path, then distortions must take place.

4.3 Long Run Dynamics under Imperfect Information

We now consider the long run dynamics of the best sustainable equilibrium. While the previous section establishes that distortions and turnover occur along the equilibrium path, it is not necessarily clear whether these persist in the long run.

More specifically, consider how dynamic incentives are provided for the policymaker along the equilibrium path. Incentive provision must take into account two constraints on the side of the policymaker. On the one hand, the policymaker must be induced to not privately rent-seek. Specifically, if the exogenous shock θ_t is high and if taxes are low, then the policymaker must be induced to not privately deviate by pretending that the shock is low and setting taxes high so as to privately raise rents for himself. Thus, if the shock is high, households can compensate the policymaker with a higher continuation value in the future relative to if the shock is low, and this can induce him to not privately deviate by setting taxes high today. This relative increase in continuation value can come in the form of higher rents and longer expected tenure in the future if the shock is high today relative to if the shock is low. Thus, providing the policymaker with incentives to not privately rent-seek implies *spreading* in the future continuation values to the policymaker.

On the other hand, the policymaker must be induced to not publicly rent-seek by setting taxes equal to the maximum today and being thrown out of power with certainty in the future. Providing such incentives to the policymaker involve *backloading* incentives to the policymaker.

In particular, the continuation value to the policymaker increases in the future in order to provide the policymaker with incentives to not expropriate today. This allows households to minimize the current level of rents paid to the policymaker today while simultaneously inducing him to not deviate from equilibrium policies. Importantly, backloading incentives to the policymaker comes together with distortions today but a reduction in distortions in the future, since a higher continuation value relaxes constraint (19) and hence enables households to invest a higher level while preserving the policymaker's incentives to not expropriate.

A natural question is whether it is the case that, in the long run, one lucky policymaker experiences a sufficiently large number of consecutive high shocks that his continuation value rises to a level where distortions disappear and he remains in power forever. From Lemma 1, it is clear that such an arrangement is sustainable; there exists an equilibrium in which, there are no distortions, transfers to the government are fixed and independent of the shock, and a permanent dictator is in place. The below proposition formally states that the economy does not converge to such an equilibrium by showing that turnover and economic distortions persist in the long run:

Proposition 5 ***(long run)*** *In the best sustainable equilibrium, turnover persists in the long run (i.e.,* $\lim_{t\to\infty} \Pr\{P_{t+k} = 1\ \forall k \geq 0\} = 0$*) and economic distortions persist in the long run (i.e.,* $\lim_{t\to\infty} \Pr\{i_{t+k} = i^*\ \forall k \geq 0\} = 0$*).*

Proof. See Appendix. ■

This proposition states that a stationary equilibrium with no distortions and a permanent dictator does not emerge in the best sustainable equilibrium. The intuition for this result is as follows. Along the equilibrium path, it is always efficient to spread future continuation values. This follows from the fact that both the policymaker and the households are risk averse, and it is efficient for them to share the risk associated with the shock θ_t by varying transfers to the government in response to θ_t. Doing so requires providing incentives for the policymaker to not privately rent-seek and then requires future continuation values to spread out (i.e., tomorrow's continuation value exceeds today's continuation value if the shock N occurs and it is below today's continuation value if the shock 1 occurs).

This incentive scheme explains why starting from any interior point $V < \overline{V}$, a policymaker experiences a positive probability of a sufficient number of consecutive low shocks so as to lead his continuation value to decline to a level where distortions emerge and where he is thrown out of power. The more subtle portion of the argument behind Proposition 5 involves why a policymaker who experiences a sufficient number of consecutive high shocks does not experience an increase in his continuation value to some absorbing state associated with the highest continuation value $\overline{V}$. The reasoning behind this is that, while it is in fact the case that some lucky policymakers do experience an increase in their continuation value to $\overline{V}$ following a long enough sequence of high shocks, the continuation value $\overline{V}$ is non-absorbing. More specifically, starting from $V = \overline{V}$, optimal policy requires the continuation value to the policymaker to decline with positive probability. The intuition for this is similar to that associated with interior $V < \overline{V}$; it is always efficient for the policymaker to share the risk associated with the shock with the households. This requires policies to remain volatile, and therefore requires citizens to continue to provide the policymaker with incentives to not privately rent-seek. This is only incentive compatible by having the policymaker experience a reduction in continuation utility in the future in the event of a low shock.

As a heuristic proof for this long run result, suppose by contradiction that it were the case that the equilibrium converged to a continuation value $V = \overline{V}$ with a stationary allocation described in Lemma 1. Moreover, for simplicity, suppose there are there two shocks θ^1 and θ^2 which each 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 $\left((1-\beta)/\beta\right)\epsilon$ 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 change in household welfare equals

$$\frac{u'(\omega)}{2}\left(\frac{v'\left(f(i^*) - i^* + \theta^1\right)}{v'\left(f(i^*) - i^* + \theta^2 - \theta^1\right)} - 1\right)\epsilon > 0, \tag{23}$$

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. Forward iteration on this argument implies that the policymaker's continuation value declines to the minimum with positive probability after a sufficiently high number of consecutive low shocks. Such a reduction in continuation value leads to economic distortions and eventual turnover. Therefore, the permanent provision of dynamic incentives to the policymaker via replacement and distortions is efficient.

There are four important points to keep in mind in interpreting the result behind Proposition 5. First, the presence of distortions and turnover in the long run does not emerge as a consequence of the non-existence of equilibria without distortions or turnover. As Lemma 1 makes clear, such equilibria exist, but Proposition 5 states that they are inefficient.

Second, risk aversion on the side of the policymaker is important for this result. If it were the case for example that the policymaker were risk neutral, then the term inside (23) 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.[23]

Third, the reasoning behind this proposition relies in large part on the presence of an upper bound on the taxes which can be paid by households. It is in fact this constraint which implies both constraints (19) and (20). In the absence (20), one could construct a sustainable stationary allocation which households consume zero and the policymaker consumes rents equal to $\omega + f(i^*) - i^* + \theta_t$ 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 third 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.[24]

Finally, our long run result is in sharp contrast to that of Acemoglu, Golosov, and Tsyvinski (2008, 2010a,b). They find that if the policymaker and the citizens have the same discount factor (as in our framework), then the policymaker has permanent tenure and distortions to production vanish in the long run. The key difference between our work and this work is that we allow for asymmetric information between the policymaker and the citizens, and this leads to the different characterization of the long run dynamics of the economy. In addition, note that in contrast to this work, distortions never emerge under full information in our model. They only emerge once private information is introduced since it generates a need for equilibrium turnover, which in turn tightens the constraint of limited commitment on

[23]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.

[24]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 in finite time.

the policymaker. Therefore, it is the interaction of the limited commitment constraint and the private information constraint which leads to distortions. We turn to this is issue in the following section.

4.4 Role of Assumption 2

In this section, we consider the extent to which our results can be generalized. In particular, we consider a situation in which we relax Assumption 2 and therefore allow $\underline{V}$, the value of being thrown out of power, to be larger.

We perform this exercise not only to better understand the limitations of our results in our particular application, but to also better connect to related models in the literature. In particular, the program defined by (14) which ignores (18) is related to the work of Acemoglu, Golosov, and Tsyvinski (2008, 2010a,b) who consider a related setting with self-interested policymakers under perfect information and who show that distortions can emerge even under full information.[25] The following proposition describes the economy in a setting in which Assumption 2 is violated in a setting with full information.[26]

Proposition 6 *In the presence of full information, the best sustainable equilibrium features:*

1. *Economic distortions (i.e., $i_t < i^*$) at date* 0,
2. *Economic distortions cease in the long run:* $\lim_{t\to\infty} \Pr\{i_{t+k} = i^* \ \forall k \geq 0\} = 1$.

Proof. See Appendix. ■

To understand this proposition, note that under full information, policymakers are never replaced along the equilibrium path. This is because their actions are perfectly observable and they need only be replaced off the equilibrium path if they deviate. Keeping them in power forever is optimal since it provides them with the strongest incentives to not expropriate. Since Assumption 2 is violated, a policymaker must be paid rents along the equilibrium path in order to be induced to not expropriate. This in turn leads to the presence of economic distortions as a means of reducing both the rents paid to the policymaker as well as the resources which he can expropriate. However, these distortions eventually vanish, and this is a consequence of *backloading*; the policymaker receives lower rents earlier in his tenure and higher rents later in his tenure. This is because rents in the future serve to relax both the current as well as the future limited commitment constraints. This implies that distortions are only required in the earlier periods when rents are low so as to prevent the policymaker from deviating. Assumption 1 implies that, eventually, the level of rents rises to a level which is high enough that distortions eventually vanish.

[25] These authors effectively assume that $\underline{V}(1-\beta) = v(0)$ which would violate Assumption 2 in our framework. The program we solve is also related to other work, such as Ray (2002) and Thomas and Worrall (1990).

[26] We additionally assume that constraint (17) never binds along the equilibrium path for expositional simplicity.

The following proposition explores the extent to which these conclusions are preserved once private information is introduced.

Proposition 7 *Under private information, the best sustainable equilibrium has the following features*

1. *There are economic distortions (i.e., $i_t < i^*$) whenever there is turnover (i.e.,* $\Pr\{P_t = 0\} > 0$*),*

2. *Economic distortions persist in the long run (i.e.,* $\lim_{t\to\infty} \Pr\{i_{t+k} = i^* \ \forall k \geq 0\} = 0$*).*

Proof. See Appendix. ■

The reasoning behind Proposition 7 is related to Propositions 4 and 5. Distortions emerge during an incumbent's first period in power since these serve to minimize his rents while simultaneously providing him with incentives to not publicly expropriate. Moreover, distortions persist in the long run because dynamic incentive provision; the policymaker's welfare declines along the equilibrium path, and eventually, investment must decline in order to induce the policymaker to not publicly expropriate.

Proposition 7 highlights the way in which the presence of long run distortions in our environment *does not depend on the size of uncertainty*. More specifically, suppose that $\theta_t = \{-\sigma, \sigma\}$ for some $\sigma > 0$, where each state occurs with probability $1/2$. Moreover, suppose that both the state and rent-seeking are privately observed by the policymaker. In this circumstance, distortions persist in the long run even for σ arbitrarily close to 0, and this is implied by Proposition 7.[27] Nevertheless, if $\sigma = 0$, then households can effectively deduce the level of rent-seeking by observation of their own consumption, so that Proposition 6 applies and distortions vanish in the long run. Therefore, our model shows how the prediction of zero long run distortions is robust to even arbitrarily small levels of uncertainty.

Note that Propositions 6 and 7 do not provide results regarding the presence of turnover. This is because if Assumption 2 is violated so that $\underline{V}$ is not sufficiently low, then it may be that replacement is too costly for society in terms of the economic distortions it entails to be used in equilibrium. More specifically, while the prospect of future replacement can be useful in the sense that it reduces the policymaker's incentives to privately rent-seek today, it also reduces the value that the policymaker places on remaining in power in the future. This in turn increases his incentives to publicly expropriate today since the value of remaining in power is not high enough relative to the value of being thrown out of power. This means that if replacement is used, high levels of economic distortions may become necessary so as to limit the current resources under the ruler's control, and these distortions can be very costly for society. Therefore, while distortions always emerge in equilibrium, it is not always the case that replacement does.

[27] Note that this point cannot be made easily in our benchmark environment which imposes Assumption 2 since Assumptions 2 and 3 cannot both simultaneously hold as σ becomes arbitrarily close to zero.

5 Simulation

In this section we explore the transitional dynamics in our model through the use of a numerical simulation. Because the constraint set represented by (16) − (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. 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

$$u(c) = c^{\sigma_u}; \quad v(x) = x^{\sigma_v}; \quad f(i) = i^{\vartheta}. \tag{24}$$

In our benchmark simulation we choose the following parameters:[28]

Table 1: Benchmark parameters for simulations

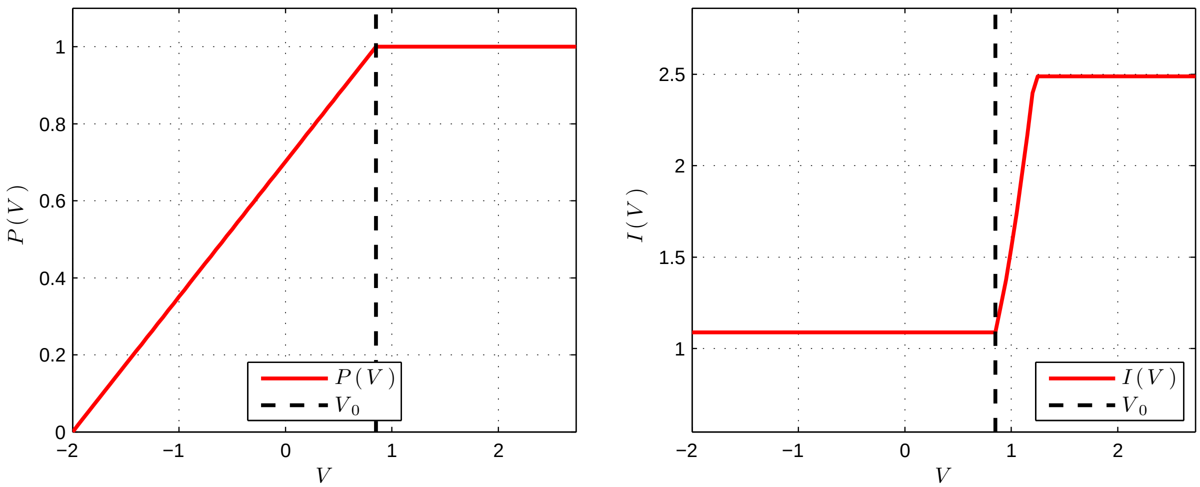
β	$\sigma_u = \sigma_v$	ϑ	ω	(θ^1, θ^2)	$\pi(\theta^1)$	$\underline{V}$
0.5	0.5	0.8	2.5	$(1.0, 1.5)$	0.5	-2

Figures 1-3 depict the policy functions conditional on the state variable V, the continuation value promised to the policymaker. Figure 1 (a) illustrates the probability of replacement as a function of the 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. 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.

Figure 1 (b) 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 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.

Figure 2 displays the level of household consumption and policymaker rents as a function of the continuation value to the policymaker. We let the subscript h and l denote the high and low shocks, respectively. Note that in this situation, the policymaker and the households share risk: both consume more during the high shock and both consume less during the low

[28] These parameters violate Assumption 2 as in section 4.4.



(a) (b)

Figure 1: Panel (a): Probability of remaining in power next period as function of V. Panel (b) Representative household's investment as function of V

shock. Clearly, the fact that the consumption of households responds to the shock implies that the level of taxes also responds to the shock. Figure 3 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 results underlying Proposition 5. If a policymaker experiences a negative economic shock, his rents in the future decline as well as his tenure. Eventually, he experiences a long enough sequence of shocks that he is necessarily replaced and the economy faces investment distortions.

6 Connection to Empirical Evidence

While the focus of our paper is on our theoretical results, we consider the extent to which the model can explain some of the empirical patterns regarding the relationship between rents, turnover, investment, and production.

First, the model suggests that policymakers are punished for negative economic shocks with shorter tenure and with lower rents. This is consistent with the evidence which suggests that policymakers are kept or replaced in response to economic shocks (e.g., Wolfers (2007), Achen and Bartels (2004), Fair (1978), Lewis-Beck (1990)). As is the case in the model, it is often argued that these shocks are beyond the control of the policymaker, so that policymakers

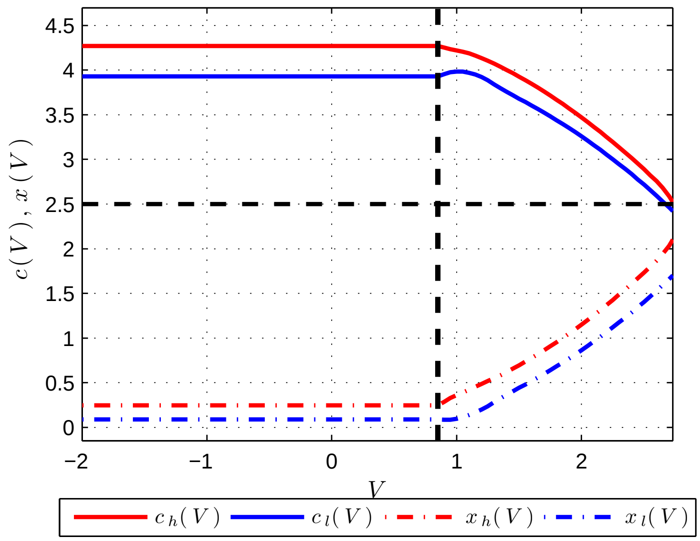


Figure 2: Representative household's consumption and government rents as function of V.

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 a whenever income increases. This is also consistent with the predictions of the model.

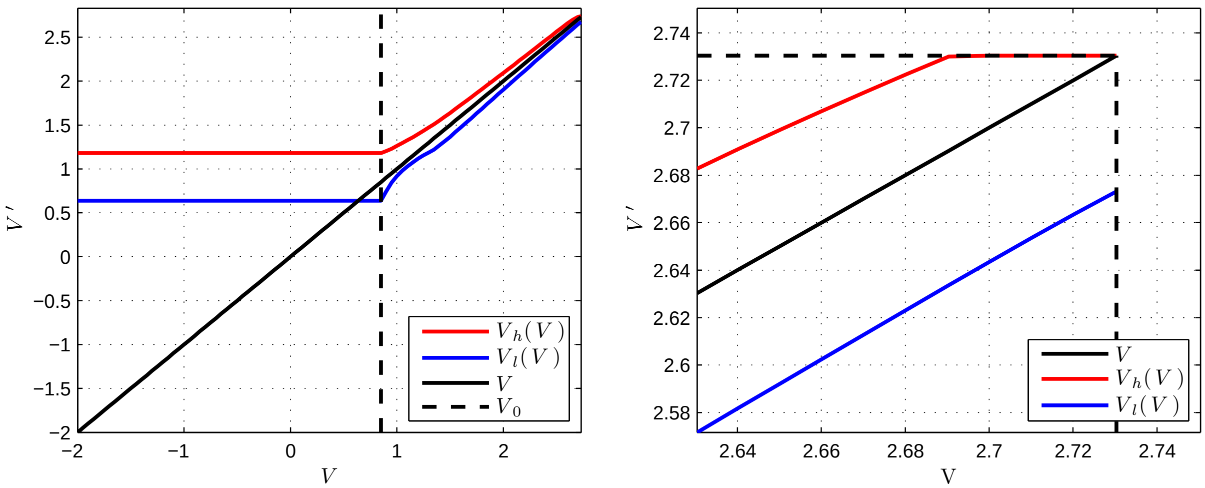
Moreover, our model also predicts history dependence in the provision of incentives to policymakers. To investigate this possibility in the data, we estimate the following equation:

$$\begin{aligned} Turnover_{it} &= \alpha NegativeGrowth_{it-1} + \beta NegativeGrowth_{it-2} \\ &+ \gamma \left(NegativeGrowth_{it-1} \times NegativeGrowth_{it-2} \right) + \eta_i + \eta_t + \epsilon_{it}. \end{aligned} \tag{25}$$

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 is not the same as in year $t-1$). $NegativeShock_{it-1}$ is a 0/1 dummy variable which takes a value of 1 if the growth rate in GDP per capita between $t-2$ and $t-1$ is below the sample mean.[29] η_i is a country fixed effect which controls for the country-level propensity for turnover and negative shocks and η_t is a time fixed effect which controls for global trends in turnover and negative shocks. ϵ_{it} is an error term. Motivated by the observation that in our model leadership transitions can occur in every period, we focus our attention on the sample of non-democracies for which leadership transitions are coded as "Irregular."[30] Given that dynamic incentives are provided in our model, one would expect

[29]Our results are robust to using the change in log GDP per capita instead of this dummy variable. Given the noise in the calculation of GDP per capita for this sample of countries, we prefer this cruder measure.

[30]Our measures of economic activity are from Heston, Summers, and Aten (2011). We combine this dataset with the Goemans, Gleditsch, and Chiozza (2009) dataset on leadership transitions. The sample excludes leaders who



(a) (b)

Figure 3: Panel (a): Continuation utility as a function of V. Panel (b) detail for values of V close to $\overline{V}$.

that $\alpha > 0$, $\beta > 0$, and $\gamma > 0$ so that the experience of negative growth has a persistent effect on tenure length. Table 2 presents different estimations of (25) and provides suggestive evidence for the history dependence in turnover. For example, Column 3 suggests that the individual effect of negative growth in either of the previous two years increases the likelihood of turnover by approximately 3.5%, and the interaction effect of two consecutive years of negative growth is almost 8%.[31]

A second prediction of our model is that investment is depressed (i.e., there are distortions) around periods of political turnover. Using the same sample as in the previous exercise, in Table 3, we explore the extent to which political turnover is associated with reduced investment. The equation we estimate is

$$\log\left(Investment/GDP\right)_{it} = \alpha Turnover_{it} + \beta Turnover_{it-1} + \beta Turnover_{it+1} + \eta_i + \eta_t + \epsilon_{it}. \quad (26)$$

$Investment/GDP$ corresponds to the investment to GDP ratio and all other terms are defined

are still in power. The transitions occur through revolts, coups, or assassinations and are in contrast to leadership transitions which occur through regular means such as elections or the natural death of leaders. We exclude from this sample transitions coded as irregular which occurred in democracies such as the assassination of John F. Kennedy, and we do so by excluding countries with a Polity composite score of 9 or 10 (Marshall and Jaggers (2004)). This provides us with an unbalanced panel of leadership transitions from 1951 to 2003.

[31] While the coefficients in this specification are not statistically significant, the F-test for all coefficient finds them to be jointly significant at the 1% level. We also considered the same specification for the entire world sample, which includes turnover through elections and natural deaths. We find that the direction of the effect is the same as in our benchmark sample, but the economic magnitude and significance of the coefficient is smaller in all cases. This is consistent with the fact that turnover is legally not feasible in many of these cases since elections are held on a constitutionally mandated schedule. Additional details are available upon request.

Table 2

Economic Shocks and Political Turnover

	(1)	(2)	(3)
Irregular Transitions	*Dependent Variable is Political Turnover*		
Negative Growth $_{t-1}$	0.0883	0.0707	0.0358
	(0.0247)	(0.0293)	(0.0342)
Negative Growth $_{t-2}$		0.0736	0.0365
		(0.0277)	(0.0280)
Negative Growth $_{t-1}$ x Negative Growth $_{t-2}$			0.0778
			(0.0507)
Growth F-test	[0.00]	[0.00]	[0.01]
Time Effects	Y	Y	Y
Country Fixed Effects	Y	Y	Y
Observations	1,078	908	908
Countries	70	69	69
R-squared	0.200	0.234	0.236

Fixed effects OLS regression in all columns with country dummies, with robust standard errors clustered by country in parentheses. Growth F-test corresponds to the p-value for the joint significance of all growth terms. Year dummies are included in all regressions. Dependent variable is political turnover. Base sample is an unbalanced panel, with annual data from 1951 to 2003 where the start date of the panel refers to the independent variable (i.e., t=1951, so t+1=1952). Panel includes countries only if they experience a leadership transition coded as "Irregular". Columns 1 excludes leaders in power for less than 365 days and columns 2-3 exclude leaders in power for less than 730 .

as in equation (25). Column 1 shows that turnover is associated with a lower investment to GDP ratio. In particular, the coefficient implies that turnover is associated with a reduction in this ratio by almost 8%. Column 2 considers the relationship between investment and turnover in the immediate past and the immediate future. We find that the contemporaneous relationship between investment and turnover is broadly unchanged. Turnover in the following year is associated with a reduction in investment relative to GDP in the current year by almost 7%, and turnover in the previous year is associated with a reduction in investment relative to GDP by almost 14%. All coefficients are statistically significant at the 1% percent level. In column 3 we add contemporaneous GDP, lagged GDP, and lagged investment to GDP ratio as additional controls. While the coefficient on turnover is diminished, it continues to be significant at the 5% level in all specifications, and it continues to imply a reduction in investment by at least 5% in all specifications. This evidence is consistent with the predictions of the model.

7 Conclusion

In this paper, we have analyzed the best sustainable equilibrium in an economy with non-benevolent policymakers who lack commitment and have private information. As in Ferejohn (1986), we show that the presence of private information creates endogenous political turnover. The key insight which emerges from this framework is that the presence of endogenous turnover creates economic distortions. Moreover, in contrast to a model with full information, our model produces long run dynamics in rents, turnover, and production.

While we focus on a production economy with a self-interested policymaker, we believe

Table 3
Political Turnover and Investment

	(1)	(2)	(3)
	Irregular Transitions		
	Dependent Variable is log(Investment/GDP)		
Turnover $_t$	-0.0794	-0.0868	-0.0587
	(0.0329)	(0.0337)	(0.0195)
Turnover $_{t-1}$		-0.137	-0.0549
		(0.0378)	(0.0236)
Turnover$_{t+1}$		-0.0694	-0.0494
		(0.0275)	(0.0202)
Turnover F-test	[0.02]	[0.00]	[0.00]
Additional Controls	N	N	Y
Time Effects	Y	Y	Y
Country Fixed Effects	Y	Y	Y
Observations	1,208	1,179	1,162
Countries	74	74	72
R-squared	0.785	0.792	0.902

Fixed effects OLS regression in all columns with country dummies, with robust standard errors clustered by country in parentheses. Turnover F-test corresponds to the p-value for the joint significance of all turnover terms. Year dummies are included in all regressions. Dependent variable is Investment/GDP ratio. Base sample is an unbalanced panel, with annual data from 1950 to 2003 where the start date of the panel refers to the independent variable (i.e., t=1950, so t+1=1951). Panel includes countries only if they experience a leadership transition coded as "Irregular". Additional controls include: log(Investment/GDP)$_{t-1}$, log(GDP)$_t$, and log(GDP)$_{t-1}$. See text for data definitions and sources.

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.[32] 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. 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

[32]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.

into account both roles for replacement by allowing for multiple types of policymakers.

References

Acemoglu, D. (2003): "Why not a political Coase theorem? Social conflict, commitment, and politics," *Journal of Comparative Economics*, 31(4), 620–652.

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*.

Acemoglu, D., S. Johnson, and J. Robinson (2005): "Institutions as a fundamental cause of long-run growth," *Handbook of economic growth*, 1, 385–472.

Acemoglu, D., and T. Verdier (2000): "The choice between market failures and corruption," *American Economic Review*, pp. 194–211.

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., M. Amador, and G. Gopinath (2009): "Investment cycles and sovereign debt overhang," *Review of Economic Studies*, 76(1), 1–31.

Albuquerque, R., and H. Hopenhayn (2004): "Optimal lending contracts and firm dynamics," *Review of Economic Studies*, 71(2), 285–315.

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*.

Banerjee, A. (1997): "A Theory of Misgovernance*," *Quarterly Journal of Economics*, 112(4), 1289–1332.

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.

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

Besley, T. (2006): *Principled agents?: The political economy of good government*. Oxford University Press, USA.

Buchanan, J., and G. Tullock (1967): *The calculus of consent: Logical foundations of constitutional democracy*, vol. 100. University of Michigan Press.

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.

Drazen, A., and E. Ilzetzki (2011): "Kosher pork," *NBER Working Papers*.

Eaton, L. (2006): "In Louisiana, Graft Inquiries are Increasing," *New York Times*, March 18th.

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.

GAO-Report (2007): "Hurricane Katrina: Ineffective FEMA Oversight of Housing Maintenance Contracts in Mississippi Resulted in Millions of Dollars of Waste and Potential Fraud," *United States Government Accountability Office*, 08-106.

——— (2010): "Oil and Gas Management: Interior's Oil and Gas Production Verification Efforts Do Not Provide Reasonable Assurance of Accurate Measurement of Production Volumes," *United States Government Accountability Office*, 10-313.

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

Heston, A., R. Summers, and B. Aten (2011): *Penn World Table Version 7.0*Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Kocherlakota, N. (1996): "Implications of efficient risk sharing without commitment," *The Review of Economic Studies*, 63(4), 595.

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.

Lipton, E. (2005): "Governor's Relative is Big Contract Winner," *New York Times*, December 7th.

Marshall, M. G., and K. Jaggers (2004): "Political Regime Characteristics and Transitions, 1800-2002," *Polity IV Project, University of Maryland*.

North, D., et al. (1981): *Structure and change in economic history*. Norton New York.

Olken, B. (2007): "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, 115(2).

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

Phelan, C. (1998): "On the long run implications of repeated moral hazard," *Journal of Economic Theory*, 79(2), 174–191.

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.

Rogoff, K., and A. Sibert (1988): "Elections and Macroeconomic Policy Cycles," *The Review of Economic Studies*, 55(1), 1.

Shleifer, A., and R. Vishny (1993): "Corruption," *The Quarterly Journal of Economics*, pp. 599–617.

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 (1988): "Self-enforcing wage contracts," *The Review of Economic Studies*, 55(4), 541.

——— (1990): "Income fluctuation and asymmetric information: An example of a repeated principal-agent problem* 1," *Journal of Economic Theory*, 51(2), 367–390.

Wolfers, J. (2007): "Are Voters Rational? Evidence from Gubernatorial Elections," Discussion paper, Working paper.

Yared, P. (2010): "Politicians, Taxes and Debt," *Review of Economic Studies*, 77(2), 806–840.

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 $i = 0$ 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 $V(q_t, z_t, \widehat{\theta})$. 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 100 percent if a household has not chosen the prescribed investment sequence up to an 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}, but otherwise all of the allocations from $t+1$ onward are unchanged (i.e., the continuation strategies are the same as if the equilibrium path replacement decisions had taken place). Moreover, following any public deviation by the representative citizen, the equilibrium allocations from $t+1$ onward are also unchanged analogously. 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 $V(q_t, z_t, \widehat{\theta})$. 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 $\underline{V}$ from $t+1$, the best public deviation maximizes immediate rents, and this is achieved with a 100 percent tax. This yields rents equal to $f\left(i_t\left(q_t, z_t\right)\right) + \theta_t$ 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.■

A.2 Proof of Lemma 2

In order to prove Lemma 2, we establish the following preliminary technical results in the below lemma. To complete this lemma, we define $C_{n,n+k}$ as follows:

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

Lemma 4 *The following is true of the solution to* (14) − (21) $\forall z$*:*

1. (19) *for* $\theta = \theta^n$ *is implied by* (19) *for* $\theta = \theta^1$ *and* (18) *for* $\theta = \theta^n$*,*
2. $x\left(\theta, z\right) - \theta$ *is weakly decreasing in* θ *and* $V'\left(\theta, z\right)$ *is weakly increasing in* θ*,*
3. (17) *for* $\theta = \theta^n$ *is implied by* (17) *for* $\theta = \theta^1$*,*
4. *If* $C_{n+1,n} \geq 0$ *and* $C_{n,n+1} \geq 0$ *for all* $n < N$*, or if* $C_{n+1,n} = 0$*, then* $C_{n+k,n} \geq 0$ *and* $C_{n,n+k} \geq 0$ *for all* n *and all* k *with* $n + k \leq N$*,*
5. *There exists a solution for which* $C_{n+1,n} = 0$ *for all* $n < N$*,*
6. (19) *for* θ^1 *is implied if* (17) *binds for* $\theta = \theta^1$*, and*
   (17) *for* θ^1 *is implied if* (19) *binds for* $\theta = \theta^1$*,*
7. (20) *does not bind for some* z *if* $V < \overline{V}$ *and* (20) *binds for all* z *if* $V = \overline{V}$*,*
8. *If* (20) *does not bind, then* $c\left(\theta, z\right) > 0$ *for all* θ*, and*
9. *If* $P\left(z\right) = 0$ *and* $V > \underline{V}$*, then* $J\left(V\right) = \overline{J}$ *and* $J'\left(V\right)$ *is differentiable with* $J'\left(V\right) = 0$*,*

Proof. Proof of part (i) Equation (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

$$v\left(x\left(\theta^n, z\right)\right) + \beta V'\left(\theta^n, z\right) \geq v(x(\theta^1, z) + \theta^n - \theta^1) - v(x(\theta^1, z)) + v(f\left(i\right) + \theta^1) + \beta \underline{V}. \tag{27}$$

The left hand side of (27) equals the left hand side of (19) for $\theta = \theta^n$. The concavity of $v\left(\cdot\right)$ implies that the right hand side of (27) weakly exceeds $v\left(f\left(i\right) + \theta^n\right) + \beta \underline{V}$ since (17) implies that $x\left(\theta^1, z\right) \leq f\left(i\right) + \theta^1$.

Proof of part (ii). 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 $v\left(\cdot\right)$ 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 $x\left(\theta, z\right) - \theta$ 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 (iii). Suppose that (17) holds for $\theta = \theta^1$, from part (ii) it holds for all θ^n.

Proof of part (iv). This is proved by induction. Suppose that $C_{n+1,n} \geq 0$ and $C_{n,n+1} \geq 0$ for all $n < N$. 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 $v\left(\cdot\right)$ implies that

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

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 $n + k < N$.

Now suppose that $C_{n+1,n} = 0$. 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

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

from concavity of $v\left(\cdot\right)$ the above implies that $x(\theta^{n+1}) - \theta^{n+1} > x(\theta^n) - \theta^n$ which contradicts point (ii).

Proof of part (v). Consider a solution to the program α for which (18) $C_{n+1} > C_n$ 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:

$$\sum_{\theta\in\Theta} \pi(\theta)\,\widehat{V}'(\theta,z) = \sum_{\theta\in\Theta} \pi(\theta)\, V'(\theta,z) \tag{28}$$

$$\widehat{V}'(\theta^{n+1},z) = \widehat{V}'(\theta^{n},z) + v(x(\theta^{n}) + \theta^{n+1} - \theta^{n}) - v(x(\theta^{n+1}))/\beta \tag{29}$$

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 (28) and the fact that (28) is also satisfied in the original allocation. From (29), it satisfies $C_{n+1,n} = 0$. Moreover, it satisfies $C_{n,n+1} \geq 0$ 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 $x(\theta^{n}) \geq x(\theta^{n+1}) - (\theta^{n+1} - \theta^{n})$ established in part (ii). From part (iv), this implies that (18) is satisfied for all θ and $\widehat{\theta}$. From part (i), we need only verify (19) for $\theta = \theta^{1}$, since (19) for other θ's are implied by the satisfaction of (18). This is implied by the fact that (28) and (29) imply that $\widehat{V}'(\theta^{1},z) \geq V'(\theta^{1},z)$. To see why this is true, note that the fact that $C_{n+1,n} \geq 0$ in the original solution implies that

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

which combined with (28) and (29) implies that $\widehat{V}'(\theta^{1},z) \geq V'(\theta^{1},z)$. Analogous arguments imply that $\widehat{V}'(\theta^{N},z) \leq V'(\theta^{N},z)$, which together with part (ii) implies that (21) is satisfied. Given (28) and (29) and the weak concavity of $J(\cdot)$, it follows that for all z,

$$\sum_{\theta\in\Theta} \pi(\theta)\, J\left(\widehat{V}'(\theta,z)\right) \geq \sum_{\theta\in\Theta} \pi(\theta)\, J\left(V'(\theta,z)\right), \tag{30}$$

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 (30), it weakly increases the welfare of the households.

Proof of part (vi). Suppose that (17) binds for $\theta = \theta^{1}$. From (21), $V'(\theta^{1},z) \geq \underline{V}$ and together with (17) which is an equality implies that (19) is satisfied for $\theta = \theta^{1}$. Suppose that (19) binds for $\theta = \theta^{1}$. Then given that from (21), $V'(\theta^{1},z) \geq \underline{V}$, it follows that (17) is implied for $\theta = \theta^{1}$.

Proof of part (vii). Suppose that (20) binds for all z if $V < \overline{V}$. 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 possible 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 exceeds $u(\omega)/(1-\beta)$, violating the fact that $J(V) = u(\omega)/(1-\beta)$ for all V.

Now suppose that $V = \overline{V}$ but that (20) does not bind for all 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 part (v), we can consider such a solution for which $C_{n+1,n} = 0$ for all n. Suppose it were the case that $C_{N-1,N} > 0$. Then, if $V'(\theta^N, z) < \overline{V}$, it would be possible to increase $V'(\theta^N, z)$ by $\epsilon > 0$ arbitrarily small while continuing to satisfy the constraints of the problem and leaving the policymaker strictly better off, violating the definition of $\overline{V}$. Analogous arguments apply if $c(\theta^N, z) > 0$ since it would be possible increase i by $\epsilon > 0$ arbitrarily small, increase $x(\theta^N, z)$ by $f(i+\epsilon) - f(i)$ for all θ and reduce $c(\theta^N, z)$ by ϵ without violating any constraints of the problem and making the policymaker strictly better off. This means that such a perturbation is not possible if $V'(\theta^N, z) = \overline{V}$ and $c(\theta^N, z) = 0$. If that is the case, then part (ii) implies that $c(\theta^n, z) = 0$ for all θ^n, which given that $C_{n+1,n} = 0$ for all n implies that $V'(\theta^n, z) = \overline{V}$ for all θ^n. But if that is the case, then this implies that households are receiving a consumption of 0 forever, which violates (20). Now suppose it were instead the case that $C_{N-1,N} = 0$ so that $V'(\theta^N, z) = V'(\theta^{N-1}, z)$ and $c'(\theta^N, z) = c'(\theta^{N-1}, z)$. Then analogous arguments to the above case would hold with respect to a perturbation on $V'(\theta^N, z)$ and $V'(\theta^{N-1}, z)$ or a perturbation on $x(\theta^N, z)$ and $x(\theta^{N-1}, z)$. Moreover, analogous arguments would rule out the situation in which $V'(\theta^N, z) = V'(\theta^{N-1}, z) = \overline{V}$ and $c'(\theta^N, z) = c'(\theta^{N-1}, z) = 0$. Forward induction on this argument on this argument implies that it cannot be that (20) does not bind.

Proof of part (viii). Suppose (20) does not bind for some z. This implies from part (vii) that $V < \overline{V}$. Moreover, given that $J(V)$ is downward sloping and weakly concave, this implies that there exists a set of sustainable allocations which can solve the sequence problem, so that the sequence problem starting from a promised value V admits well-defined Lagrange multipliers on all constraints. Given the Inada condition on $u(\cdot)$, this implies that $c(\theta, z)$ is interior.

Proof of part (ix). Note that $V_0 > \underline{V}$. This is because from (19),

$$V_0 \geq \sum_{\theta \in \Theta} \pi(\theta) v(f(i) + \theta) + \beta \underline{V} > v(0) + \beta \underline{V} \geq \underline{V}.$$

Now suppose that $P(z) = 0$ for some z and $J(V) \neq \overline{J}$. This would mean that $V > V_0$ since otherwise $J(V) = \overline{J}$. We can show that this is not possible. This would imply that $J(V) < \overline{J}$ by definition of $\overline{J}$. Now consider the solution α given V and define $q = \int_0^1 P(z)\,dz$ and

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

It is clear that since $V > V_0$, that satisfaction of (15) requires $V_q > V > V_0$. It follows from (14) and the fact that $J(\cdot)$ is weakly concave that

$$J(V) = (1-q)\,\overline{J} + qJ\left(V_q\right) < \overline{J}. \tag{31}$$

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

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

which contradicts (31) since $\widetilde{q} < q$. This establishes that $V \in (\underline{V}, V_0]$ and $J(V) = \overline{J}$.

Now suppose that $V > \underline{V}$ and $P(z) = 0$ for some z. The above reasoning implies that $V \in (\underline{V}, V_0]$. In order to prove that $J(V)$ is differentiable with $J'(V) = 0$, it is sufficient to show that $V < V_0$, since the above reasoning implies that $J(V+\epsilon) = \overline{J}$ for $|\epsilon| > 0$ arbitrarily small. To see why $V < V_0$, suppose by contradiction that $V = V_0$ and that $P(z) = 0$ for some z. Define q and V_q as above. This would imply that $V_q > V_0$ where analogous reasoning to the above arguments would imply that $J(V_0) = (1-q)\,\overline{J} + qJ\left(V_q\right) = \overline{J}$ which means that $J\left(V_q\right) = \overline{J}$. This would however violate the definition of V_0 since V_0 represents the highest welfare which can be provided for the policymaker conditional on providing households with their highest welfare $\overline{J}$. ∎

We now can prove Lemma 2.

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

$$J\left(V^\kappa\right) \geq \kappa\left(V'\right) + (1-\kappa)\,J\left(V''\right), \tag{32}$$

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:

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

where we have taken into account 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

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

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

We now prove that $J(V)$ is strictly decreasing in V if $V \in [V_0, \overline{V}]$. This is because if this were not the case, then given the weak concavity of $J(\cdot)$, this would imply that $J(V) = \overline{J}$ for all V. However, if this is true, then this would violate the definition of V_0, since V_0 which must represent the highest continuation value that the policymaker can receive conditional on the households receiving their highest continuation welfare $\overline{J}$.

Proof of part (iii). We now prove the continuous differentiability of $J(V)$ for $V \in (\underline{V}, \overline{V})$. From part (ix) of Lemma 4, we only need to consider continuation values for which $P(z) = 1$ for all z, since otherwise $J(V) = \overline{J}$ and $J'(V) = 0$. We now consider these cases and use Lemma 1 of Benveniste Scheinkman (1979). According to this result, if there exists a function $Q(V + \epsilon)$ for $\epsilon \gtreqless 0$ which is differentiable, weakly concave, and satisfies

$$Q(V + \epsilon) \leq J(V + \epsilon) \tag{33}$$

for arbitrarily small values $|\epsilon|$ where (33) is an equality if $\epsilon = 0$, then $J(V)$ is continuously differentiable at V. Given the solution α for a given V which generates welfare $J(V)$, we now establish the existence of such a function $Q(\cdot)$ by constructing a perturbed solution $\widehat{\alpha}(\epsilon)$

which satisfies $\widehat{\alpha}(0) = \alpha$ as well as (16) − (21) for a given promised value $V + \epsilon$, where the welfare associated with this sustainable solution is defined as $Q(V + \epsilon)$ and must necessarily satisfy (33). Note that by part (v) of Lemma 4, we can consider an original solution α where $C_{n+1,n} = 0$. There are two cases to consider.

Case 1. By parts (vii) and (viii) of Lemma 4, it must be the case that in the original solution α that for some z, $c(\theta, z) > 0$ for all θ. Suppose that for all such z, there exists a set of z's for which $i(z) > 0$ and $x(\theta, z) > 0$ for all θ. Let $\widetilde{Z}$ correspond to such z's for which $i(z) > 0$ and $c(\theta, z) > 0$ and $x(\theta, z)$ for all θ, and let $q_{\widetilde{Z}} = \Pr(z \in \widetilde{Z})$. Given ϵ, let $\widehat{\alpha}(\epsilon)$ be identical to α, with the exception that

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

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

$$\widehat{c}(\theta^{n},z,\epsilon) + \widehat{x}(\theta^{n},z,\epsilon) = \widehat{i}(z,\epsilon) + \theta^{n} \ \forall\theta^{n} \tag{34}$$

$$\sum_{\theta^{n}\in\Theta} \pi(\theta^{n})\, v(x(\theta^{n},z)) + \epsilon/q_{\widetilde{Z}} = \sum_{\theta^{n}\in\Theta} \pi(\theta^{n})\, v(\widehat{x}(\theta^{n},z,\epsilon)) \tag{35}$$

$$v(\widehat{x}(\theta^{n+1},z,\epsilon)) - v(\widehat{x}(\theta^{n},z,\epsilon) + \theta^{n+1} - \theta^{n}) = v(x(\theta^{n+1},z)) - v(x(\theta^{n},z) + \theta^{n+1} - \theta^{n}), \forall\theta^{n}, n < N \tag{36}$$

$$v\left(\widehat{x}\left(\theta^{1},z,\epsilon\right)\right) - v(f(\widehat{i}(z,\epsilon)) + \theta^{1}) = v\left(x\left(\theta^{1},z\right)\right) - v\left(f(i(z)) + \theta^{1}\right) \tag{37}$$

for a given ϵ. To see that $\xi(z,\epsilon)$ exists, note that (35) − (37) corresponds to five equality constraints and $\left\{i(z,\epsilon), \{c(\theta,z,\epsilon), x(\theta,z,\epsilon), \}_{\theta\in\Theta}\right\}$ corresponds to five non-negative unknowns for sufficiently low ϵ. 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. The allocation $\widehat{\alpha}(\epsilon)$ implies a households welfare $Q(V+\epsilon)$ which is a continuously differentiable function of ϵ with $Q(V) = J(V)$.

We are left to 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 solution to the program, implying that (33) must hold. (34) guarantees that $\widehat{\alpha}(\epsilon)$ satisfies (7) and (35) guarantees that $\widehat{\alpha}(\epsilon)$ satisfies promise keeping. Now let us check that (18) is satisfied. To do this, we appeal to part (iv) of Lemma 4 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, (36) 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 (36) 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 $\widehat{\alpha}(\epsilon)$. We now verify that (17) and (19) are satisfied if $\theta = \theta^1$ under $\widehat{\alpha}(\epsilon)$. This is implied by part (vi) of Lemma 4, by the fact that (17) and (19) are satisfied if $\theta = \theta^1$ under α, and by (37). Parts (i) and (iii) of Lemma 4 then imply that (17) and (19) are satisfied for all θ. Since (20) is slack under α, then it is also slack under $\widehat{\alpha}(\epsilon)$. 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 (33) holds which implies that $J(V)$ is differentiable.

Case 2. Suppose that for all z for which $c(\theta, z) > 0$ for all θ, there does not exist a set of z's for which $i(z) > 0$ and $x(\theta, z) > 0$ for all θ. Let $\widetilde{Z}$ correspond to the z's for which $c(\theta, z) > 0$ for all θ, and let $q_{\widetilde{Z}} = \Pr\left(z \in \widetilde{Z}\right)$. We will prove that in this case, $J'(V) = 0$. To do this, construct $Q(V+\epsilon)$ for $\epsilon > 0$ as in case 1, where it can easily be verified from (35) $-$ (37) that $\xi(z, \epsilon)$ exists since $\left\{i(z,\epsilon), \{c(\theta,z,\epsilon), x(\theta,z,\epsilon),\}_{\theta\in\Theta}\right\}$ are all non-negative for sufficiently low ϵ. Every element of $\widehat{\alpha}(\epsilon)$ satisfies (16) $-$ (21) for a given promised value $V+\epsilon$ by analogous reasoning as in case 1. Furthermore, it follows by implicit differentiation given the Inada conditions on $v(\cdot)$ and $f(\cdot)$ that

$$\lim_{\epsilon>0,\epsilon\to0} \frac{Q(V+\epsilon) - Q(V)}{\epsilon} = 0 \leq \lim_{\epsilon>0,\epsilon\to0} \frac{J(V+\epsilon) - J(V)}{\epsilon}, \tag{38}$$

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 (38) binds with equality. Since $J(V)$ is weakly decreasing and weakly concave, it follows that

$$0 \geq \lim_{\epsilon>0,\epsilon\to0} \frac{J(V) - J(V-\epsilon)}{\epsilon} \geq \lim_{\epsilon>0,\epsilon\to0} \frac{J(V+\epsilon) - J(V)}{\epsilon} = 0,$$

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

B Proofs of Section 4.2

B.1 Proof of Proposition 2

Consider the solution to (14) which ignores constraint (18) so that there is no private information. Let us solve the relaxed problem which additionally ignores (19) and we will verify that (19) is satisfied. First order conditions with respect to investment imply that $i_t = i^*$ for all t and the optimal level of rents satisfies $x_t = 0$. This maximizes the welfare of households. The replacement rule which additionally maximizes the welfare of the period zero incumbent features $P_t = 1$ for all t. Assumption 2 implies that (19) is satisfied for all θ_t.■

B.2 Proof of Lemma 3

By the same arguments as those of Proposition 2, if there exists an allocation with $i_t = i^*$ and $x_t = 0$ for all t, then this allocation is the solution to the program since it achieves the highest welfare for the households. We now show how to sustain such equilibrium determining the sequence of P_t's. We let $q(\theta_{t-1}) = \Pr\{P_t = 1|\theta_{t-1}\}$, so that the replacement probability in a given date t depends only on the realization of the shock at date $t-1$. In the constructed equilibrium, $q(\theta_{t-1})$ must satisfy the following system of equations:

$$\begin{aligned} V(\theta^N) &= v(0) + \beta \sum_{n=1}^{N} \pi(\theta^n) V(\theta^n) \\ v(0) + \beta V(\theta^n) &= v(\theta^n - \theta^{n-1}) + \beta V(\theta^{n-1}) \text{ for } n > 1, \text{ and} \\ V(\theta^n) &= q(\theta^n) V(\theta^N) + (1 - q(\theta^n)) \underline{V} \text{ for } n \geq 1. \end{aligned} \tag{39}$$

This system of equations yields a unique solution for $V(\theta^n)$ and $q(\theta^n)$. Moreover, for such a solution, it can be easily verified that $q(\theta^n)$ is strictly increasing in θ^n with $q(\theta^N) = 1$. Therefore, feasibility of the sequence of P_t's which satisfies this system requires that $q(\theta^1) \geq 0$. This is guaranteed to be true if there exists a solution to this system with $V(\theta^1) \geq \underline{V}$. By some algebra, the system implies that $V(\theta^N)$ satisfies

$$V(\theta^N) = \frac{v(0) - \sum_{n=2}^{N} \left(v\left(\theta^n - \theta^{n-1}\right) - v(0)\right) \sum_{k=1}^{n-1} \pi\left(\theta^k\right)}{1-\beta},$$

with $V(\theta^N)$ and $V(\theta^1)$ related by the following equation:

$$V(\theta^N) = v(0) + \beta V(\theta^1) + \sum_{n=2}^{N} (v(\theta^n - \theta^{n-1}) - v(0)) \sum_{k=n}^{N} \pi(\theta^k),$$

and this implies that $v(0) + \beta V(\theta^1)$ equals the left hand side of (22). It can thus be verified given the violation of Assumption 3 that $V(\theta^1) \geq \underline{V}$. Let us first verify that (18) is satisfied. This is guaranteed to be true by (39) and part (iv) of Lemma 4. Finally, the violation of Assumption 3 guarantees that (19) holds if $\theta_t = \theta^1$, which implies that (19) holds for all θ_t by the arguments of part (i) of Lemma 4.

In any equilibrium where $x_t = 0$ for all t it must be that $\Pr\{P_{t+k} = 1 \forall k \geq 0\} = 0$ for all t. Suppose not, then there exist a positive measure of paths so that the policymaker remains in power forever and $V(q_t, z_t, \theta_t) = V((q_t, z_t, \widehat{\theta})) = v(0) / (1-\beta)$ in (18). However this implies that (18) is violated if $\theta_t > \theta^1$.■

B.3 Proof of Proposition 3

To show that replacement must occur, consider the solution for $V = V_0$. 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. By part (iv) of Lemma 4, we need only consider the local constraints for (18), let $\pi(\theta^{n+1})\,\phi(\theta^{n+1}, \theta^n, z)\,dz$ and $\pi(\theta^n)\,\phi(\theta^n, \theta^{n+1}, z)\,dz$ correspond to the Lagrange multipliers on the downward and upward incentive compatibility constraint, 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 $n=N$. Let $\beta\pi\left(\theta^n\right)\underline{\mu}\left(\theta^n,z\right)dz$ and $\beta\pi\left(\theta^n\right)\overline{\mu}\left(\theta^n,z\right)dz$, and $\pi\left(\theta^n\right)\upsilon\left(\theta^n,z\right)dz$ correspond to the Lagrange multipliers on the constraints that $V'\left(\theta^n,z\right)\geq\underline{V}$, $V'\left(\theta^n,z\right)\leq\overline{V}$, and $x\left(\theta^n,z\right)\geq 0$, respectively. The Inada conditions guarantee that the non-negativity constraints on $c\left(\theta^n,z\right)$ and $i\left(z\right)$ can be ignored. First order conditions yield:

$$u'\left(c\left(\theta^n,z\right)\right)\left(P(z)+\eta\left(z\right)\right)=\nu\left(\theta^n,z\right), \tag{40}$$

$$\left\{\begin{array}{c}\left(\begin{array}{c}P(z)\lambda+\phi\left(\theta^n,\theta^{n+1},z\right)\\ +\phi\left(\theta^n,\theta^{n-1},z\right)+\psi\left(\theta^n,z\right)\end{array}\right)v'\left(x\left(\theta^n,z\right)\right)\\ -\phi\left(\theta^{n-1},\theta^n,z\right)v'\left(x\left(\theta^n,z\right)+\theta^{n-1}-\theta^n\right)\frac{\pi\left(\theta^{n-1}\right)}{\pi\left(\theta^n\right)}\\ -\phi\left(\theta^{n+1},\theta^n,z\right)v'\left(x\left(\theta^n,z\right)+\theta^{n+1}-\theta^n\right)\frac{\pi\left(\theta^{n+1}\right)}{\pi\left(\theta^n\right)}\end{array}\right\}=\nu\left(\theta^n,z\right)+\kappa\left(\theta^n,z\right)-\upsilon\left(\theta^n,z\right) \tag{41}$$

$$f'\left(i\left(z\right)\right)-1=\frac{\sum_{n=1}^{N}\pi\left(\theta^n\right)\left(-\kappa\left(\theta^n,z\right)f'\left(i\left(z\right)\right)+\psi\left(\theta^n,z\right)v'\left(f\left(i\left(z\right)\right)+\theta^n\right)f'\left(i\left(z\right)\right)\right)}{\sum_{n=1}^{N}\pi\left(\theta\right)\nu\left(\theta^n,z\right)} \tag{42}$$

$$J'\left(V'\left(\theta^n,z\right)\right)\left(P(z)+\eta\left(z\right)\right)=\left\{\begin{array}{c}-P(z)\lambda-\phi\left(\theta^n,\theta^{n-1},z\right)-\phi\left(\theta^n,\theta^{n+1},z\right)\\ +\phi\left(\theta^{n-1},\theta^n,z\right)\frac{\pi\left(\theta^{n-1}\right)}{\pi\left(\theta^n\right)}+\phi\left(\theta^{n+1},\theta^n,z\right)\frac{\pi\left(\theta^{n+1}\right)}{\pi\left(\theta^n\right)}\\ -\psi\left(\theta^n,z\right)-\underline{\mu}\left(\theta^n,z\right)+\overline{\mu}\left(\theta^n,z\right)\end{array}\right\} \tag{43}$$

and the Envelope condition yields:

$$J'\left(V\right)=-\lambda. \tag{44}$$

From Lemma 2 part (ii) we have that $J'\left(V_0\right)=0$. Suppose by contradiction that replacement never takes place. Since rents weakly exceed 0, this implies that $V_0\geq v\left(0\right)/\left(1-\beta\right)$ and $V'\left(\theta,z\right)\geq v\left(0\right)/\left(1-\beta\right)$ so that $\underline{\mu}\left(\theta^n,z\right)=0$ for all n. Given Assumption 2, this means that (19) does not bind so that $\psi\left(\theta^n,z\right)=0$ for all n. Note that given that $J\left(V_0\right)>u\left(\omega\right)/\left(1-\beta\right)$, it follows that $\eta\left(z\right)=0$ for all z, since otherwise it would be possible to make households strictly better off while satisfying all of the constraints of the program by never choosing allocations for which (20) binds. From (43) and (44), this implies that given z,

$$\sum_{n=1}^{N}\pi\left(\theta^n\right)J'\left(V'\left(\theta^n,z\right)\right)=J'\left(V_0\right)+\sum_{n=1}^{N}\pi\left(\theta^n\right)\overline{\mu}\left(\theta^n,z\right)=\sum_{n=1}^{N}\pi\left(\theta^n\right)\overline{\mu}\left(\theta^n,z\right). \tag{45}$$

Given that $J\left(\cdot\right)$ is weakly decreasing it follows that the left hand side of equation (45) is weakly negative. This implies that for all n, $\overline{\mu}\left(\theta^n,z\right)=0$ and $J'\left(V'\left(\theta^n,z\right)\right)=0$. From part (ii) of Lemma 2 implies that

$$V'\left(\theta^n,z\right)\leq V_0,\qquad\forall\ n. \tag{46}$$

Therefore, from (43) 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 (43) for all 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 $\phi\left(\theta^N, \theta^{N-1}, z\right) > 0$. Equation (18) implies that

$$x(\theta^{N-1}) = x(\theta^N) - (\theta^N - \theta^{N-1}) < x(\theta^N). \tag{47}$$

Now consider (41) for $n = N$ given that $\phi\left(\theta^N, \theta^{N-1}, z\right) > 0$:

$$0 > \phi(\theta^N, \theta^{N-1}, z)[v'(x(\theta^N, z)) - v'(x(\theta^N, z) - (\theta^N - \theta^{N-1}))] = \nu(\theta^N, z) + \kappa(\theta^N, z) - v(\theta^N, z) \tag{48}$$

From (40), it must be that $\nu(\theta^N, z) > 0$ and constraint (17) implies that $\kappa(\theta^N, z) \geq 0$, which means that for (48) to hold, it must be that $v(\theta^N, z) > 0$ so that $x(\theta^N, z) = 0$. However, this implies from (47) 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 (47) and (48) 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 (41) 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 $v\left(\theta^n, z\right) > 0$ and $x\left(\theta^n, z\right) = 0$ for all n. Therefore,

$$V_0 = v\left(0\right) + \int_0^1 \left[\left(\sum_{n=1}^{N} \pi\left(\theta^n\right) \beta V'\left(\theta^n, z\right)\right)\right] dz \leq v\left(0\right) + \beta V_0,$$

where we have appealed to (46), and this implies that $V_0 \leq v\left(0\right) / \left(1 - \beta\right)$. Given that the policymaker is never replaced, this can only be true if $V = v\left(0\right) / \left(1 - \beta\right)$ so that rents are equal to 0 in every period with $V'\left(\theta^n, z\right) = v\left(0\right) / \left(1 - \beta\right)$. However, if $x\left(\theta^n, z\right) = 0$ and $V'\left(\theta^n, z\right) = v\left(0\right) / \left(1 - \beta\right)$ for all n and z, then (18) is violated, leading to a contradiction. ■

B.3.1 Proof of Proposition 4

From part (ix) of Lemma 4, we only need to show that this is true for $V = V_0$. This is because if $\Pr\left\{P_t = 0\right\} > 0$, then it implies that $V_t \in [\underline{V}, V_0]$, and therefore that the continuation value of the policymaker who is in power following the replacement decision is V_0. To show that it must be that $i_0 < i^*$. There are several cases to consider.

Case 1. Conditional on z, suppose that $\psi\left(\theta^n, z\right) > 0$ for some θ^n. From part (i) of Lemma 4, this would only be the case for $n = 1$. Moreover, from parts (iii) and (vi) of Lemma 4, constraint (17) is made redundant for all n and can be ignored since (19) binds with an equality, so that $\kappa(\theta^n, z) = 0$. Given that $\nu\left(\theta^n, z\right) > 0$ from (40), this means that the right hand side of (42) is positive so that $i\left(z\right) < i^*$.

Case 2. Conditional on z, suppose that $\psi\left(\theta^n,z\right)=0$ for all n. (43) and (44) imply that

$$\sum_{n=1}^{N}\pi\left(\theta\right)J'\left(V'\left(\theta^n,z\right)\right)=J'\left(V_0\right)+\sum_{n=1}^{N}\pi\left(\theta^n\right)\left(-\underline{\mu}\left(\theta^n,z\right)+\overline{\mu}\left(\theta^n,z\right)\right). \tag{49}$$

There are now two cases two consider.

Case 2a. Suppose that $\underline{\mu}\left(\theta^n,z\right)=0$ for all n. Then given that $J'\left(V'\left(\theta^n,z\right)\right)\leq 0$ and $J'\left(V_0\right)=0$, the above equation then implies that $\overline{\mu}\left(\theta^n,z\right)=0$ and that $J'\left(V'\left(\theta^n,z\right)\right)=0$ for all n so that (46) applies. It follows from the same arguments as those in the proof of Proposition 3 that $\phi\left(\theta^n,\theta^{n+1},z\right)=\phi\left(\theta^{n+1},\theta^n,z\right)=0$ for all $n<N$ and that $x\left(\theta^n,z\right)=0$ for all n. Therefore, (18) implies that for all $n>1$: $v\left(0\right)+\beta V'\left(\theta^n,z\right)\geq v(\theta^n-\theta^{n-1})+\beta V'(\theta^{n-1})$. Analogous arguments those used in the proof of Lemma 3 then implies that the value of $v\left(0\right)+\beta V'\left(\theta^1,z\right)$ cannot exceed the left hand side of (22). But if this is the case, (19) under $\theta=\theta^1$ is violated if $i\left(z\right)\geq i^*$. Therefore, it is necessary that $i\left(z\right)<i^*$.

Case 2b. Suppose that $\underline{\mu}\left(\theta^n,z\right)>0$ for some n. From part (ii) of Lemma 4, it is the case that $V'\left(\theta^n,z\right)$ is weakly increasing in n, so that this implies that $\underline{\mu}\left(\theta^1,z\right)>0$ so that $V'\left(\theta^1,z\right)=\underline{V}$. Therefore, $J'\left(V'\left(\theta^1,z\right)\right)=0$ and from (43), this implies that $\frac{\pi(\theta^2)}{\pi(\theta^1)}\phi\left(\theta^2,\theta^1,z\right)-\phi\left(\theta^1,\theta^2,z\right)=\underline{\mu}\left(\theta^1,z\right)>0$, where we have used the fact that $\lambda=\psi\left(\theta^1,z\right)=\overline{\mu}\left(\theta^1,z\right)=0$. Moreover, given (17) and (19), the fact that $V'\left(\theta^1,z\right)=\underline{V}$ implies that $x\left(\theta^1,z\right)=f\left(i\left(z\right)\right)+\theta^1$. Suppose that $x\left(\theta^1,z\right)=0$. Given Assumption 1, this would imply that $i\left(z\right)<i^*$, since Assumption 1 guarantees that $f\left(i\left(z\right)\right)+\theta^1>0$ for $i\left(z\right)\geq i^*$. Therefore, we are left to consider the case for which $x\left(\theta^1,z\right)>0$ so that $v\left(\theta^1,z\right)=0$. Substitution into (41) for $n=1$ yields:

$$\phi(\theta^1,\theta^2,z)v'(f(i\left(z\right))+\theta^1)-\frac{\pi(\theta^2)}{\pi(\theta^1)}\phi(\theta^2,\theta^1,z)v'(f\left(i\left(z\right)\right)+\theta^2)\geq 0,$$

where we have used the fact that $\nu\left(\theta^1,z\right)\geq 0$ and $\kappa\left(\theta^1,z\right)\geq 0$. In order for this condition to hold, it is necessary that $\phi\left(\theta^1,\theta^2,z\right)>0$, which given that $\frac{\pi(\theta^2)}{\pi(\theta^1)}\phi\left(\theta^2,\theta^1,z\right)-\phi\left(\theta^1,\theta^2,z\right)\geq 0$ implies that $\phi\left(\theta^2,\theta^1,z\right)>0$. From (18), this implies that $x\left(\theta^2,z\right)=x\left(\theta^1,z\right)+\theta^2-\theta^1$, which means that $x\left(\theta^2,z\right)=f\left(i\left(z\right)\right)+\theta^2>0$. Moreover, this implies that $V'\left(\theta^2,z\right)=V'\left(\theta^1,z\right)=\underline{V}$ so that $J'\left(V'\left(\theta^2,z\right)\right)=0$. Repeating this argument as in the previous case implies that $x\left(\theta^n,z\right)=f\left(i\left(z\right)\right)+\theta^n$ and $V'\left(\theta^n,z\right)=\underline{V}$ for all n. This means that $J'\left(V'\left(\theta^n,z\right)\right)=0$ and $\overline{\mu}\left(\theta^n,z\right)=0$ for all n. However, given that $J'\left(V_0\right)=0$ and $\underline{\mu}\left(\theta^n,z\right)\geq 0$ for all n with $\underline{\mu}\left(\theta^1,z\right)>0$, this violates (49).■

C Proofs of Section 4.3

C.1 Proof of Proposition 5

In order to prove this proposition, we first prove two important lemmas which establish conditions under which the continuation values to the policymaker decline.

Lemma 5 *Suppose that the solution to* (14) − (21) *for a given* $V \in (\underline{V}, \overline{V})$ *has the following properties: the elements of* α *do not depend on the value of* z*,* $P(z) = 1$ *for all* z*, and* (19) *is a strict inequality for all* θ*. Then it must be that the solution admits* $J'(V'(\theta^1, z)) > J'(V)$ *for all* z*.*

Proof. Suppose this were not the case so that $J'(V'(\theta^1, z)) \leq J'(V)$ for all z. Part (ii) of Lemma 4 together with the weak concavity of $J(V)$ then implies that $J'(V'(\theta^n, z)) \leq J'(V)$ for all n and z. From the arguments used in the proof of part (v) of Lemma 4, 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 $C_{n+1} = C_n$ for all $n < N$. Note that this perturbation weakly increases the value of $V'(\theta^1, z)$ so that it remains the case that $J'(V'(\theta^1, z)) \leq J'(V)$ for all 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):

$$J(V) = \max_{\alpha} \left\{ \int_0^1 \left[\begin{array}{c} \sum_{n=1}^{N} \pi(\theta^n) \left(u(c(\theta^n, z)) + \beta J(V'(\theta^n, z)) \right) \\ +\lambda \left(\sum_{n=1}^{N} \pi(\theta^n) \left(v(x(\theta^n, z)) + \beta V'(\theta^n, z) \right) \right) \end{array} \right] dz \right\} \tag{50}$$

s.t.

$$c(\theta^n, z) + x(\theta^n, z) = \omega - i(z) + f(i(z)) + \theta^n \ \forall n, z \tag{51}$$

$$x(\theta^n, z) \leq f(i(z)) + \theta^n \ \forall n, z \tag{52}$$

$$x(\theta^{n+1}, z) \leq x(\theta^n, z) + \theta^{n+1} - \theta^n \ \forall n < N, z \tag{53}$$

$$v(x(\theta^{n+1}, z)) + \beta V'(\theta^{n+1}, z) = v(x(\theta^n, z) + \theta^{n+1} - \theta^n) + \beta V'(\theta^n, z) \ \ \forall n < N, z \tag{54}$$

$$\sum_{n=1}^{N} \pi(\theta^n) u\left(c(\theta^n, z) + \beta J\left(V'(\theta^n, z)\right)\right) \geq u(\omega) / (1 - \beta) \ \ \forall z, \tag{55}$$

$$\text{and } V'(\theta^n, z) \in \left[\underline{V}, \overline{V}\right] \ \ \forall n, z. \tag{56}$$

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 (53) and (54) by using part (iv) of Lemma 4. Given the results in Propositions 3 and 4, we only need consider the case for which $\lambda > 0$. Define Lagrange multipliers $\pi(\theta) \phi(\theta^{n+1}, \theta^n, z)\, dz$, $\beta\pi(\theta^n)\, \underline{\mu}(\theta^n, z)\, dz$, and $\beta\pi(\theta^n)\, \overline{\mu}(\theta^n, z)\, dz$ as in the proof of Proposition 3. First order conditions with respect to $V'(\theta^n, z)$ yield:

$$J'(V'(\theta^n, z))(1 + \eta(z)) = -\lambda - \phi\left(\theta^n, \theta^{n-1}, z\right) + \phi\left(\theta^{n+1}, \theta^n, z\right) + \overline{\mu}(\theta^n, z), \tag{57}$$

where we have taken into account that the fact that $V'(\theta^n, z) \geq V$ implies that $V'(\theta^n, z) > \underline{V}$ so that $\underline{\mu}(\theta^n, z) = 0$. The envelope condition yields: $J'(V) = -\lambda$. From (57), 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 $\phi\left(\theta^N,\theta^{N-1},z\right) \leq \phi\left(\theta^2,\theta^1,z\right) \leq 0$. (57) 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 $\phi\left(\theta^{n+1},\theta^n,z\right) = 0$ for all $n<N$. Therefore, constraint (54) can be ignored.

Let us assume and later verify that constraint (53) can also be ignored. Then, first order conditions with respect to $c(\theta^n,z)$ and $x(\theta^n,z)$ together with (51) imply that

$$\lambda v'(x(\theta^n,z)) \geq u'(\omega - i(z) + f(i(z)) + \theta^n - x(\theta^n,z)), \tag{58}$$

which is a strict inequality only if $x(\theta^n,z) = f(i(z)) + \theta^n$. It is easy to verify that constraint (53) is satisfied under such a solution. This is because the value of $x(\theta^n,z)$ for which (58) 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 (58) 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 (53) 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 (54) is ignored, it follows that these values are uniquely defined conditional on λ. Given that $J'(V'(\theta^n,z)) = -\lambda$ for all n, it follows by forward iteration on the recursive program that $c(\theta^n,z)$, $x(\theta^n,z)$, and $i(z)$ are independent of time which means that $V'(\theta^n,z)$ does not vary across n. Given that $C_{n+1} = C_n$ 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 $x(\theta^n,z) = f(i(z)) + \theta^n$ for all n. If that is the case, then the welfare of households given V equals $u(\omega - i(z))/(1-\beta)$, which means that for (55) to be satisfied, it must be the case that $i(z)=0$. However, if that is the case, it violates the first order condition with respect to $i(z)$ in (42), where we have used the fact that $\psi(\theta^n,z)=0$ since (19) is a strict inequality for all θ^n. ■

Lemma 6 *If $V = \overline{V}$, then the solution to (14) − (21) admits $V'\left(\theta^1,z\right) < V$ for all z.*

Proof. Suppose that $V=\overline{V}$. Part (ix) of Lemma 4 implies that the solution admits $P(z)=1$ for all z and part (vii) of Lemma 4 implies that (20) binds for all z and the arguments in the proof of part (vii) of Lemma 4 imply that the policymaker achieves a continuation value of $\overline{V}$ for all z.

Suppose it were the case that for some z, $V'\left(\theta^1,z\right) \geq V$, which given part (ii) of Lemma 4 as well as (21) implies that $V'(\theta^n,z) = \overline{V}$ for all n. This implies that $J(V'(\theta^n,z)) = u(\omega)/(1-\beta)$ by part (vii) of Lemma 4 which from (20) means that $c(\theta^n,z) = \omega$ for all n. Satisfaction of (18) in this situation implies that $x(\theta^n,z) = x\left(\theta^{n+1},z\right) + \theta^{n+1} - \theta^n$ for all

$n < N$. Together with the fact that $\omega + f(i) - i \leq \omega + f(i^*) - i^*$ and that $c(\theta^n, z) = \omega$ for all n, this means that $x(\theta^n, z) \leq f(i^*) - i^* + \theta^n$. We now show that the previous relationship must hold with equality for all n. Suppose it were the case that $x(\theta^1, z) < f(i^*) - i^* + \theta^1$. Then this would imply that $x(\theta^n, z) < f(i^*) - i^* + \theta^n$ so that

$$\overline{V} = \sum_{n=1}^{N} \pi(\theta^n) v(x(\theta^n, z)) / (1-\beta) < \sum_{n=1}^{N} \pi(\theta^n) v(f(i^*) - i^* + \theta^n) / (1-\beta). \tag{59}$$

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 (59). This however contradicts the fact that $\overline{V}$ corresponds to the highest sustainable welfare for the policymaker. Therefore, it is necessary that $x(\theta^n, z) = f(i^*) - i^* + \theta^n$ for all 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 lower bound on J' 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 $J'(\overline{V}) = \lim_{\epsilon \to 0^+} \left(J(\overline{V}) - J(\overline{V} - \epsilon)\right) / \epsilon$. Then it must be that

$$J'(\overline{V}) \leq -u'(\omega) / \left(\sum_{n=1}^{N} \pi(\theta^n) v'(f(i^*) - i^* + \theta^n) \right). \tag{60}$$

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

$$\epsilon = \sum_{n=1}^{N} \pi(\theta^n) \left(v(f(i^*) - i^* + \theta^n) - v(f(i^*) - i^* + \theta^n - \varepsilon(\epsilon)) \right). \tag{61}$$

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 the statement of the claim.

Given the bound in (60), 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

$$dV^n(\epsilon) = \frac{1}{\beta}\left(\sum_{n=1}^{N}\pi(\theta^n)v(f(i^*)-i^*+\theta^n+\epsilon)-v(f(i^*)-i^*+\theta^n)\right) = \\ = -\sum_{n=1}^{N}\pi(\theta^n)v'(f(i^*)-i^*+\theta^n)\epsilon \tag{62}$$

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

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

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:

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

Substituting (62) and (60) into the above implies that $\pi(\theta^N)u'(\omega)\lim_{\epsilon\to 0^+}\frac{dx^N(\epsilon)}{\epsilon}\le 0$, which is a contradiction since dx^N is strictly positive. ■

Using these two lemma, we will now prove Proposition 5 through a sequence of steps:

Step 1. Suppose that starting from some date t, replacement ceases to occur so that $P_{t+k}=1\ \forall k\ge 0$. This implies that V_{t+k}, the continuation value to a policymaker at a given date $t+k$ prior to the realization of z_{t+k}, weakly exceeds V_0. This is because if were below V_0, then application of the result in Proposition 3 implies that replacement occurs with positive probability and there are distortions.

Step 2. In this situation, for any given V_{t+k}, $\Pr\{V_{t+k+1}<V_{t+k}\}\ge\pi\left(\theta^1\right)$. To see why, note first that by Assumption 2, constraint (19) is a strict inequality from t onward since the absence of replacement means that $V_{t+k}\ge v(0)/(1-\beta)$ for all k. This means that starting from any given $V_{t+k}=V$, it is the case that the solution to (14) $-$ (21) admits $\psi(\theta^n,z)=0$ for all n and all z so that (19) can be ignored. Therefore, if $V_{t+k}=\overline{V}$, then by Lemma 6 $V_{t+k+1}<V_t$ if $\theta_{t+k}=\theta^1$, so that the statement is true. If instead $V_{t+k}=V\in\left(\underline{V},\overline{V}\right)$, it then follows from the arguments in Lemma 5 that for every z: $V'(\theta^1,z)<V$, where this is because for every z, $J'\left(V'\left(\theta^1,z\right)\right)>J'(V)$. Therefore, in this case, $V_{t+k+1}<V_t$ if $\theta_{t+k}=\theta^1$.

Step 3. Let $\underline{V}'$ correspond to the infinum of all of the possible realizations of V_{t+k} for all $k>0$, so that $\Pr\left\{V_{t+k}\in\left[\underline{V}',\overline{V}\right]\right\}=1$, where it is clear from Step 1 that $\underline{V}'>V_0$. It follows that for any $V'\in\left[\underline{V}',\overline{V}\right]$ and for any V with $V_{t+k}=V$, there exists a l large enough such that if state 1 is repeated l times consecutively, then $V_{t+k+l} < V'$. To see why this is true, let $g(V)$ correspond to the highest realization of $V'(\theta^1, z)$ 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, $V_{t+k+l} \leq g(V_{t+k+l-1}) < V_{t+k+l-1}$, where we have used step 2. Suppose by contradiction that $\lim_{l\to\infty} V_{t+k+l} = V'' \geq V'$. This implies that $\lim_{l\to\infty} g(V_{t+k+l-1}) \geq V'' \geq V'$. 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.

Step 4. For each $V_{t+k} \in [\underline{V}', \overline{V}]$, let $l(V_{t+k})$ correspond to the number of consecutive realizations of state 1 required for $V_{t+k+l} < \underline{V}'$. By step 3 for each $V_{t+k} \in [\underline{V}', \overline{V}]$, $l(V_{t+k})$ exists and is finite. Let L correspond to the maximum such $l(V_{t+k})$. It follows that for $k \geq L$,

$$\Pr\left\{V_{t+k} < \underline{V}' | V_t \in [\underline{V}', \overline{V}]\right\} \geq \Pr\{\theta_{t+k} = ... = \theta_{t+k-L} = \theta^1\} = [\pi(\theta^1)]^L.$$

Therefore, it is not possible $\Pr\left\{V_{t+k} \in [\underline{V}', \overline{V}]\right\} = 1$ for all k. Therefore, there is replacement and distortions in the long run.■

D Proofs of Section 4.4

D.1 Proof of Proposition 6

The following preliminary lemma applies to an environment which ignores (18)

Lemma 7 *In the case of full information, if (20) binds for some z, then $V'(\theta^n, z) = \overline{V}$ for all n,*

Proof. Suppose that (20) binds conditional on z so that households receive a continuation value equal to $u(\omega)/(1-\beta)$. Then in this situation, the policymaker must receive a continuation value conditional on z which cannot exceed the solution to the following program:

$$V^{\max} = \max_{\{x(\theta^n,z)\}_{n=1}^N} \sum_{n=1}^N \pi(\theta^n) v(x(\theta^n, z)) / (1-\beta)$$

$$c(\theta^n, z) + x(\theta^n, z) = \omega + f(i^*) - i^* + \theta^n \ \forall n,$$

$$\sum_{n=1}^N \pi(\theta^n) u(c(\theta^n, z)) dz / (1-\beta) = u(\omega) / (1-\beta).$$

This follows from the fact that $V^{\max}$ corresponds to the highest continuation value which can be achieved by the policymaker subject only to the constraint that (20) is respected.[33] Clearly, $V^{\max} \geq \overline{V}$. It can be verified that $V^{\max} = \overline{V}$, and this is because the stationary allocation which solves $V^{\max}$ from Assumption 1, satisfies all sustainability constraints. Therefore, efficiency

[33]The program which solves $V^{\max}$ takes into account that i^* corresponds to the efficient level of investment.

requires that if (20) binds conditional on z, the policymaker achieves a welfare of $V^{\max}$, which means that $V'(\theta^n, z) = \overline{V}$ for all n. ■

Back to the proof of proposition 6:

Proof of part (i). Consider the solution to the program which ignores (18). First order conditions and the Envelope condition are the same as (40) − (44) with $\phi\left(\theta^n, \theta^{n+1}, z\right) = \phi\left(\theta^n, \theta^{n-1}, z\right) = 0$. Since we consider the equilibrium for the new incumbent with $V = V_0$, it is the case that $\lambda = 0$. We assume that (17) can be ignored so that $\kappa(\theta^n, z) = 0$ for all θ^n, z. From (42), this means that $i(z) \le i^*$, and if it is the case that $i(z) = i^*$, this means that $\psi(\theta^n, z) = 0$ for all θ^n. If this is the case, then the above reasoning implies that $x(\theta^n, z) = 0$ for all θ^n. Moreover, from (43), (44) and the concavity of $J(V)$, it follows that $V'(\theta^n, z) \le V_0$. Therefore, $V_0 \le v(0) + \beta V_0$, which implies that $V_0 \le v(0) / (1-\beta)$. However, if this it the case, then constraint (19) is violated and this follows from the fact that Assumption 2 is violated. Hence $i(z) < i^*$.

Proof of part (ii). (43) and (44) imply that

$$J'\left(V'(\theta^n, z)\right)(1 + \eta(z)) = J'(V) - \psi(\theta^n, z) - \underline{\mu}(\theta^n, z) + \overline{\mu}(\theta^n, z). \tag{63}$$

We first establish that $V'(\theta^n, z) \ge V$, and there are two cases to consider. First, suppose that constraint (20) does not bind so that $\eta(z) = 0$. Then (63) implies that $J'(V'(\theta^n, z)) \le J'(V)$ which given the concavity of $J(V)$, this means that $V'(\theta^n, z) \ge V$. Second, suppose that constraint (20) does bind. Then Lemma 7 implies that $V'(\theta^n, z) = \overline{V} \ge V$. Therefore, $\{V_t\}_{t=0}^{\infty}$ corresponds to a weakly increasing stochastic sequence which is bounded from above by $\overline{V}$. Therefore, it converges. However, for this to be true, it is necessary from (63) that $\psi(\theta^n, z) = 0$. From (42), this means that $\lim_{t\to\infty} \Pr\{i_{t+k} = i^* \ \forall k \ge 0\} = 1$).■

D.2 Proof of Proposition 7

Proof of part (i). This follows from the same arguments as those used in the proof of Proposition 4 which do not rely on Assumption 2.

Proof of part (ii). The same arguments as those used in the proof of Proposition 5 can be used where we take into account that if there were no economic distortions from date t onward, it would be necessary for the continuation value to the policymaker V_{t+k} strictly exceed V_0 for all $k \ge 0$, where this follows from the fact that if $V_{t+k} = V_0$, distortions would emerge and this follows from part (i).■