Politico-Economic Inequality and the Comovement of Government Purchases*

Rüdiger Bachmann[a,†] Jinhui H. Bai[b,‡]

[a] RWTH Aachen University, Templergraben 64, Rm. 318, 52062 Aachen, Germany.

[b] Georgetown University, Department of Economics, Washington, DC 20057-1036, United States.

September 12, 2012

Abstract

This paper explores the implications of economic and political inequality for the comovement of government purchases with macroeconomic fluctuations. We set up and compute a heterogeneous-agent neoclassical growth model, where households value government purchases which are financed by income taxes. A key feature of the model is a wealth bias in the political aggregation process. When calibrated to U.S. wealth inequality and exposed to aggregate productivity shocks, such a model is able to generate weaker positive comovement of government purchases than models with no political wealth bias. The wealth bias that matches the cross-sectional campaign contribution distribution by income is consistent with the mild positive comovement of government purchases in the aggregate data. We thus provide an empirically relevant example where economic and political heterogeneity matter for aggregate dynamics.

JEL Codes: E30, E32, E60, E62, H30.

Keywords: government purchases, wealth bias, economic inequality, political inequality, heterogeneous agents, politico-economic equilibrium.

*We thank Marina Azzimonti and Davide Debortoli for their discussions. We are grateful to seminar participants at Boston University, the Chicago FED, the Cleveland FED, City University of Hong Kong, Georgetown University, the Konstanz Seminar on Monetary Theory and Policy, Mannheim University, University of Michigan, the 2010 Midwest Macro Meeting, Université de Montreal, National University of Singapore, 2010 NBER-SI-PEPF, the 2010 New York/Philadelphia Fed Quantitative Macro Workshop, 2010 SED (Montreal), University of Virginia, Yale, as well as Christopher House, Mark Huggett, Gianluca Violante and two anonymous referees for their comments. Any remaining errors are our own.

†Corresponding author. Phone: +49-241-8096203. Fax: +49-241-8092649. E-mail: ruediger.bachmann@rwth-aachen.de. Also: NBER (United States), CESifo (Germany), and ifo (Germany).

‡E-mail: jb543@georgetown.edu.

1 Introduction

This paper provides a quantitative theory of the implications of economic and political inequality for the dynamic comovement of government purchases with aggregate economic activity. We define government purchases as "government expenditures on consumption and investment goods" as in the NIPA accounts, excluding transfers and interest payments. In many economic environments, where government purchases are valued as a normal good, desired government purchases increase in aggregate productivity. Aggregate productivity impacts all economic agents in the same direction and will therefore, in isolation, lead to a strong positive comovement of government purchases with aggregate output. Government purchases are no different than private consumption expenditures in this respect. Unlike the decentralized private consumption choice, however, the choice of government purchases typically involves an aggregation of individual preferences. Economic and economically-based political inequality may thus be important for aggregate dynamics.

The mechanism that we explore in this paper follows from two properties of standard models: individually desired government purchases, a normal good, increase in individual wealth. Therefore, if the political aggregation mechanism is such that in good economic times agents with lower individual wealth gain relative political influence compared to normal times, then the effect of individual wealth on desired government purchases might dampen the comovement of government purchases and output. We will refer to this effect as *decoupling*. The second feature is negative comovement of wealth inequality with output, which, if combined with a pro-wealth bias in the political decision making process, indeed has agents with lower individual wealth gain relatively more political influence in periods of high economic activity.

We document for U.S. post-war data that while government purchases comove positively with aggregate economic activity, which is consistent with the normality property, they are also robustly the component of aggregate demand for which comovement is weakest. This holds for the aggregate as well as for most disaggregated components of government purchases. Also, political participation data from the American National Election Studies (ANES) is consistent with the relative political influence of lower income percentiles mildly comoving with the aggregate economy, whereas the opposite is the case for higher income percentiles. Recently, the literature has provided further evidence that is consistent with our mechanism: historically, and perhaps against conventional wisdom, U.S. states with larger inequality have leaned towards higher public good provision (Boustan et al., 2010). Notice a similar effect behind the partial decoupling of government purchases from economic activity in our mechanism: periods of strong economic activity are times of reduced wealth inequality and therefore social preferences are shifted towards less provision of public goods. It is important to stress that this is a statement about government purchases, not transfers.

It is well known that income and wealth inequality are an important feature of the U.S. economy (see Diaz-Gimenez et al., 1997, as well as Heathcote et al., 2010, for a comprehensive documentation of economic heterogeneity). A different literature provides direct evidence of wealth bias in political participation and argues that political inequality due to wealth bias is crucial in understanding the long-run *steady state* relationship between economic inequality and public policies (see Rosenstone and Hansen, 1993, Benabou, 2000, Bartels, 2008, and Campante, 2011). In this paper, we ask whether the joint presence of economic and political inequality is also important for understanding certain aspects of the *time series* dynamics of public policies.

As a first step, we illustrate normality of government purchases in a simple static model. In a second step, we ask whether the described mechanism with the interacting normality property of government purchases and negatively comoving political inequality can also generate mild positive comovement of government purchases with output and private consumption in a fully dynamic stochastic general equilibrium model that features quantitatively realistic economic and political inequality. We set up a stochastic neoclassical growth model with incomplete markets, where the equilibrium wealth distribution has a realistic concentration. Political inequality is linked to wealth inequality through the presence of a wealth bias in the political decision making process. We use a "political power"-weighted social welfare function to endogenously determine public policy.

Otherwise, the model is essentially a merger of the representative household model in Klein et al. (2008) that features endogenous public policy, and the heterogeneous agent model in Krusell and Smith (1998). This means we add to Klein et al. (2008) three features: aggregate uncertainty in the form of aggregate productivity shocks, economic heterogeneity in the form of persistent idiosyncratic shocks to labor productivity as well as – in the spirit of Krusell and Smith (1997 and 1998) – discount factors, and the parametric specification of political wealth bias from Benabou (2000). Our model features a government that cannot commit ex ante to a path of government purchases, but takes into account future streams of government purchases and how they depend on current decisions. The solution concept for the game between successive governments is the Markov-perfect equilibrium. Government purchases are financed by income taxes. Like Klein et al. (2008), we abstract from government debt and transfers.

We find that the contemporaneous correlation between output and government purchases is a declining function of the wealth bias in the political system. The degree of wealth bias that under a "One-Dollar-Of-Contribution-One-Vote" interpretation of the political system makes the model match the campaign contribution shares by income percentiles reported in the Cooperative Congressional Election Survey (CCES), a recent further development of the ANES, also produces a contemporaneous correlation between output and government purchases that is broadly in line with the data. Either a lack of political inequality or a lack of quantitatively realistic wealth inequality, the extreme case of which is a representative agent version of our model, means that output and government purchases are too strongly correlated, compared to the data. Thus one contribution of our paper is to compute and analyze successive extensions of the Klein et al. (2008) model of time-consistent government policy with aggregate uncertainty and economic and political heterogeneity. Models that lack either type of inequality feature essentially approximate aggregation in the sense of Krusell and Smith (1998) and behave very similarly to the representative agent case. Current government purchases in such models are mainly determined by the intertemporal trade-off between private and government consumption today and tomorrow, the wealth effect from aggregate productivity. By contrast, a very unequal wealth distribution, especially at the upper tails, intensifies disagreement about the optimal level of government purchases that, if funneled through political inequality, can cause a partial decoupling of aggregate government purchases and output. As in the data, in the model the political weight of the wealth-poorest comoves positively and that of the wealth-richest comoves negatively with the aggregate economy. To be clear: we do not mean to say that ours is the only conceivable mechanism that could explain the mild positive comovement of government purchases.[1] But in standard economic environments it is a consequence of quantitatively realistic economic and political inequality.

Related Literature

Besides the intrinsic link to the literature on wealth heterogeneity and political wealth bias, our paper is most closely related to the literature on endogenous government purchases in dynamic environments. This is a quantitative macro literature that starts from an otherwise standard neoclassical growth model and uses a Markov-perfect equilibrium concept to endogenize government purchases (see Klein et al., 2008, Debortoli and Nunes, 2010, and Azzimonti, 2011). This literature has so far focussed mostly on long-run steady state analyses, and has also abstracted from economic or political heterogeneity. Recently, Barseghyan et al. (2010) and Azzimonti et al. (2010), building on earlier work in Battaglini and Coate (2008), have developed a general framework to characterize fiscal policies under legislative bargaining and with aggregate uncertainty. This literature does not have the neoclassical growth model set up in its economic part. For instance, they use a utility function that is linear in private consumption, they abstract from capital accumulation and wealth inequality which is the central focus of our framework. While this literature is much richer than our paper in some dimensions in that it deals with multidimensional government policies, these simplifications might limit its

[1] Debortoli and Nunes (2010) as well as Bachmann and Bai (2011), using a representative agent model, introduce an additional aggregate shock in the political system and find a lower contemporaneous correlation between government purchases and output than in models with one aggregate shock.

usefulness for quantitative analysis. We complement their work by studying the dynamics of government purchases in a probabilistic voting and a majority voting environment.

In addition, our paper is also related to three other strands of the literature. Methodologically, we build on the framework developed in the literature on dynamic political economy, e.g., Krusell et al. (1997), Krusell and Rios-Rull (1999), Hassler et al. (2003), Hassler et al. (2005), Corbae et al. (2009), Bai and Lagunoff (2011a), Song et al. (2011) as well as Song (2012). Secondly, this paper complements the literature on procyclical fiscal policy in developing countries that focuses on international aspects such as sovereign borrowing constraints and dysfunctional democracy (see Alesina et al., 2008, Ilzetzki and Vegh, 2008, as well as Ilzetzki, 2011, for an overview). Thirdly, our paper is related to the literature with fiscal policy in incomplete markets, such as Heathcote (2005) and Gomes et al. (2008).

The remainder of the paper proceeds as follows: section two sets up the economic environment and, with the help of a simple model, discusses the main forces behind decoupling. Section three describes the equilibrium concept, the political aggregation mechanism, computation and calibration of the quantitative model. Section four presents the results from numerical simulations and some robustness checks. Various details are relegated to appendices.

2 A Simple Static Model

In this section, we describe the main mechanism by which economic and political inequality influence the dynamics of government purchases in the fully-fledged quantitative and dynamic model in Section 3. We show in a simple static model that in a standard economic environment desired government purchases of an arbitrary economic agent are an increasing function of aggregate productivity and individual wealth. The response to aggregate productivity makes government purchases positively correlated with output and private consumption. If the political aggregation mechanism is such that in periods of high economic activity wealth-poorer agents gain political influence compared to normal times, then the effect of individual wealth on desired government purchases might dampen the comovement of government purchases and output. Countercyclical wealth inequality coupled with a pro-wealth bias in the political decision about government purchases would lead to such a political aggregation mechanism.

To gain intuition we study the (hypothetical) decision problem of an agent who values private consumption, c, and government purchases, G. The agent is endowed with $\tilde{l}$ units of time, labor productivity, ϵ, and an initial capital stock, k, which depreciates at rate δ. The agent receives factor income from competitive capital rental and labor markets, subject to a linear income tax, at rate $\tau(G)$, which is a function of government purchases.

$$\max_{c,G} U(c,G) \equiv \theta u_1(c) + (1-\theta)\, u_2(G)\,,\ \text{such that } c = (1-\delta)\, k + (1-\tau(G))\left(w\tilde{l}\epsilon + rk\right). \tag{1}$$

$0 < \theta < 1$, and $u_1(c)$ as well as $u_2(G)$ are increasing, strictly concave and differentiable functions. Notice that we write this problem as if the agent were to decide about both private consumption and government purchases. This will in general not be the case. Instead, G will be decided in a political aggregation mechanism, through a constitution (see Sections 3.2 and 3.3). However, this hypothetical decision problem allows us to introduce the concept of privately desired government purchases, $\hat{G}$.

To close the model, we assume that the economy is populated by a continuum of such agents with potentially different labor productivities and capital stock endowments, that there exists an aggregate Cobb-Douglas constant-returns-to-scale technology to produce output and that the government has to balance its budget:

$$Y = zK^{\alpha}L^{1-\alpha},\ \text{where } K = \int_0^1 k_i\, di \text{ and } L = \tilde{l}\int_0^1 \epsilon_i\, di. \tag{2}$$

$$\tau(K,L,z,G) = \frac{G}{zK^{\alpha}L^{1-\alpha}}. \tag{3}$$

z denotes aggregate productivity which is known in the static model, but, in the dynamic versions of the model, will be the only source of aggregate uncertainty. Competitive factor markets guarantee the usual factor price conditions: $w(K,L,z) = (1-\alpha)\, z\, (K/L)^{\alpha}$ and $r(K,L,z) = \alpha z\, (K/L)^{\alpha-1}$.

A re-write of the budget constraint illustrates the sources of conflicting policy preferences. Plugging the tax function $\tau(K,L,z,G)$ into the budget constraint and re-arranging terms, yields:

$$c + p(k,\epsilon;K,L)\, G = (1-\delta)\, k + p(k,\epsilon;K,L)\, zK^{\alpha}L^{1-\alpha}, \tag{4}$$

where

$$p(k,\epsilon;K,L) \equiv \frac{w(K,L,z)\,\tilde{l}\epsilon + r(K,L,z)\, k}{zK^{\alpha}L^{1-\alpha}} = (1-\alpha)\frac{\tilde{l}\epsilon}{L} + \alpha\frac{k}{K} \tag{5}$$

can be viewed as the individual-specific relative price of G for a household with characteristic (k,ϵ), measured in units of private consumption. From this perspective, the decision problem falls into the framework of classical Consumer Demand Theory, where the agents face a given relative price $p(k,\epsilon;K,L)$ and endowment $\left\{(1-\delta)k, zK^{\alpha}L^{1-\alpha}\right\}$ of two commodities, c and G. The relative price does not depend on aggregate productivity, because z proportionally increases the income of every household and hence brings no change in relative income.

The effect of a change in z and k on the private demand for G, denoted by $\widehat{G}(k,\epsilon;K,L,z)$, can now be readily analyzed and summarized in the following[2]

Proposition $\widehat{G}(k,\epsilon;K,L,z)$ is strictly increasing in z. If $u_1(c)$ has the CRRA form, $\frac{c^\gamma}{\gamma}$, $\gamma \leq 0$ and $\delta \leq 1$, with one inequality holding strictly, then $\widehat{G}(k,\epsilon;K,L,z)$ is strictly increasing in k. In the case of linear consumption felicity, $\gamma = 1$, $\widehat{G}(k,\epsilon;K,L,z)$ is strictly decreasing in k, while it is independent of z.

The intuition behind this result is that with $\gamma = 1$ there is no income effect, so an increase in individual capital holdings through an increase in the individual-specific relative price of government purchases will only lead to a substitution effect. The lower γ, the stronger the income effect that makes agents want to demand more G. The proposition is important because it predicts for parameters that are standard in quantitative macroeconomics that wealth-richer agents demand more government purchases.

The proposition highlights two channels for the dynamics of government purchases. A higher state of aggregate productivity tends to increase G, which in isolation makes G comove with aggregate economic activity. This productivity-induced wealth effect does not require heterogeneity of agents. It is present even if all agents are the same. But heterogeneity in capital endowments may generate an additional channel if wealth inequality changes over time and the political system exhibits a wealth bias in the aggregation of preferences over government purchases. For instance, with a pro-wealth bias, a decrease in wealth inequality operating through a decrease in political inequality might move the policy outcome towards the preference of poorer people, hence – assuming $\widehat{G}(k,\epsilon;K,L,z)$ increases in k – towards a lower level of G. Other things equal, the higher the wealth bias, the larger is this effect. Compared to a representative agent model with only the productivity-induced wealth effect present, wealth inequality that negatively comoves with aggregate economic activity would then dampen the positive comovement of G with macroeconomic quantities. It is ultimately a quantitative question, whether these additional forces through heterogeneity are present in a fully-fledged dynamic and realistically calibrated model, which justifies our use of numerical methods.

3 The Quantitative Model

In this section, we build on the model from Section 2 and use numerical methods to study a calibrated dynamic stochastic general equilibrium model with heterogeneous agents and quantitatively realistic economic and political inequality. We first add what is necessary to specify

[2] The proof can be found in Appendix A.

the economic environment in this quantitative model. Then we define the equilibrium concept with endogenous public policy. The ensuing subsection describes the political aggregation mechanism. We finish with a discussion of the computation and calibration of the model.

3.1 The Dynamic Economic Environment

Section 2 specified the main ingredients of a standard heterogeneous household stochastic growth model, as in Krusell and Smith (1998). In the dynamic setting, agents discount the future at rate β. We specify the felicity function (1) in log-log form:

$$U(c, G) = \theta \log(c) + (1-\theta)\log(G). \tag{6}$$

The budget constraint now reads, suppressing for ease of notation the dependence of w, r and τ on aggregate variables:

$$c + k' = (1-\delta)k + (1-\tau)(w\tilde{l}\epsilon + rk). \tag{7}$$

Finally, there is also a borrowing constraint $k' \in [\underline{k}, +\infty)$.

These infinitely lived households face persistent idiosyncratic shocks to their labor efficiency, ϵ. This leads to idiosyncratic labor income risk which households partially insure using capital as the only asset. Idiosyncratic labor efficiency shocks are a standard assumption in the incomplete markets literature (see, e.g., Huggett, 1993, and Aiyagari, 1994), because they give rise to a non-degenerate wealth distribution. In order to generate quantitatively realistic wealth inequality, we – in the baseline case – again follow Krusell and Smith (1998) and assume that households face persistent idiosyncratic shocks to their discount factor. We will show that matching the inequality in the U.S. wealth distribution is crucial for understanding the dynamics of government purchases in our model. We assume that the two sources of heterogeneity, ϵ and β, evolve according to discrete Markov chains, independently of each other and across agents.[3] Similarly, z, i.e. aggregate productivity, evolves according to a discrete Markov chain, which is independent from the two Markov processes that govern the idiosyncratic stochastic environment.

3.2 Dynamic Equilibrium with Endogenous Public Policy

In choosing government purchases, G, the government faces – on top of the balanced budget requirement – two institutional constraints. First, the government chooses G under social

[3] Since ϵ is uncorrelated across households, and by the continuum assumption, L is a constant. We therefore suppress it as an argument in the real wage, the real interest rate and the tax function.

choice institutions. Some examples of social choice institutions are the utilitarian social planner and a majority voting system, although we are going to study more general mechanisms in the next subsection that allow for a wealth-bias. For the purpose of defining an equilibrium, however, we present the constraint as an abstract social preference aggregator, $W\left(\{J_i\}_{i\in[0,1]}\right)$, which maps the preferences of each household, J_i, to the equilibrium choice. J_i denotes in net present value terms the indirect utility function of each household over alternative policy proposals. It is formally defined below.

Second, the government cannot commit to a stream of future policies. Without a commitment device, it is well known that the commitment equilibrium in our environment is not time-consistent. Time consistency requires imposing a subgame-perfect restriction with successive governments and the households as game players. Following Krusell and Rios-Rull (1999) and Klein et al. (2008), we focus on a subclass of subgame-perfect equilibrium with Markov strategies, i.e., Markov-perfect Equilibrium (MPE). Adapted to our heterogeneous agent environment, the aggregate state variables consist of aggregate productivity, z, and the joint distribution over the pair $\left(k_i, \epsilon_i, \beta_i\right)$, denoted by Γ. With the choice of these state variables, the MPE is defined in terms of continuation value functions and best response functions under a one-shot deviation. Loosely speaking, MPE is achieved if these objects satisfy standard requirements of Recursive Competitive Equilibrium (RCE) on the equilibrium path, and best-response consistency on the off-equilibrium path. The formal definition follows.

Definition 1 *A Markov-perfect Equilibrium for the economy is a set of functions, including a government policy function* $G = \Psi(\Gamma, z)$*, a distribution transition function* $\Gamma' = H(\Gamma, z, G)$*, an equilibrium continuation value function* $v\left(k, \epsilon, \beta, \Gamma, z; \Psi, H\right)$*, a best-response value function* $J\left(k, \epsilon, \beta, \Gamma, z, G; \Psi, H\right)$ *and a best-response decision rule* $k' = h\left(k, \epsilon, \beta, \Gamma, z, G; \Psi, H\right)$*, such that*
(a) For any given G*, the functions* $J\left(k, \epsilon, \beta, \Gamma, z, G; \Psi, H\right)$*,* $v\left(k, \epsilon, \beta, \Gamma, z; \Psi, H\right)$ *and* $h\left(k, \epsilon, \beta, \Gamma, z, G; \Psi, H\right)$ *solve the household's problem*

$$
\begin{aligned}
J\left(k, \epsilon, \beta, \Gamma, z, G; \Psi, H\right) \quad = \quad & \max_{\{c, k'\}} \left\{ U(c, G) + \beta E\left[v\left(k', \epsilon', \beta', \Gamma', z'; \Psi, H\right) | \epsilon, z \right] \right\} \\
& s.t. \\
& c \geq 0, k' \geq \underline{k}, \\
& c + k' = (1-\delta)\, k + \left(1 - \tau(K, z, G)\right)\left(w(K, z)\, \tilde{l} \epsilon + r(K, z)\, k\right), \\
& \Gamma' = H(\Gamma, z, G).
\end{aligned}
$$

In addition, $v\left(k, \epsilon, \beta, \Gamma, z; \Psi, H\right) = J\left(k, \epsilon, \beta, \Gamma, z, \Psi(\Gamma, z); \Psi, H\right)$.
(b) $H(\Gamma, z, G)$ *is implied by* $h\left(k, \epsilon, \beta, \Gamma, z, G; \Psi, H\right)$ *and the exogenous stochastic processes.*
(c) $\Psi(\Gamma, z)$ *is a result of social choice, i.e.,* $\Psi(\Gamma, z) = W\left(\left\{J\left(k_i, \epsilon_i, \beta_i, \Gamma, z, G; \Psi, H\right)\right\}_{i\in[0,1]}\right)$.

The first part of the equilibrium definition says that h is the best response of the household to an arbitrary change in current G when the future follows the equilibrium path, a so-called one-shot deviation best response. J denotes the corresponding value function. In addition, the best-response value function should coincide with the equilibrium continuation function when evaluated at the equilibrium policy $G = \Psi(\Gamma, z)$. The second part requires that the evolution of the aggregate distribution, $H(\Gamma, z, G)$, is generated by the households' best responses for any given G and the exogenous stochastic processes. This reflects rational expectations on the household side. On the equilibrium path, this requirement reduces to the familiar consistency restriction in RCE. In addition, MPE imposes the same requirement for off-equilibrium paths. The third part imposes the constraint from the social choice institution.

3.3 The Social Choice Mechanism

In our baseline model, the public choice mechanism is defined as

$$\Psi(\Gamma, z) = \arg\max_{G} \left\{ \int (k^{+})^{\chi} \, J\left(k, \epsilon, \beta, \Gamma, z, G; \Psi, H\right) d\Gamma \right\}, \tag{8}$$

where $\chi \in \mathbb{R}$ is a given institutional parameter reflecting characteristics of the political process and $k^{+} \equiv \max[0, k]$.[4] The government chooses public policy so as to maximize a weighted social welfare function, with weights dependent on the wealth of the households. The weighting function, $(k^{+})^{\chi}$, is meant to be a flexible parametric form to capture wealth bias in the political process. If $\chi = 0$, every household is treated equally, which leads to the familiar egalitarian social welfare function. A positive (negative) value of χ implies a pro-wealth (anti-wealth) bias in the political process, since a larger (smaller) weight is assigned to a household with a higher positive wealth. As the absolute value of χ increases, the degree of wealth bias becomes larger. Intuitively, a higher wealth bias increases the responsiveness of political weights to wealth inequality, hence implies a higher political inequality. In fact, this intuition can be proved formally when political inequality is measured in terms of the Lorenz curve (see Bai and Lagunoff, 2011b, for a formal proof and a general characterization of related weighting functions).

There are different interpretations of our baseline public choice mechanism. From a normative perspective, the weighted social welfare function approach can be viewed as a social planner's problem. There is also a micro-founded, positive interpretation, where this social welfare function corresponds to a political process in a probabilistic voting environment (see Lindbeck and Weibull, 1987). Under this interpretation the weighting function, $(k^{+})^{\chi}$, can be attributed to a pro-wealth vote allocation in a weighted voting system.

[4] See Benabou (2000) and Bai and Lagunoff (2011b). Using k^{+} is a convenient way to avoid negative weights in the welfare function. We also experimented with a weighting function that depends on total available resources, $(k - \underline{k})^{\chi}$, and found that the results do not depend on the precise functional form.

Appendix B provides an alternative interpretation of our baseline economy as a richer environment with permanent *ex ante* heterogeneity in average individual labor productivity levels and preferences for government purchases of the following form: permanently more productive households are also those agents that have a lower preference weight for public goods in their felicity function (6). We also argue in this appendix that as long as these household properties are permanent or at least only slowly changing over time, the decoupling mechanism is unaffected.

3.4 Computation

Whether the partial decoupling mechanism outlined in Section 2 is economically significant, is a quantitative question. We thus use numerical methods to characterize and analyze the Markov-perfect equilibria of our economy. In computing the equilibria, we face the usual practical challenges introduced by the heterogeneous-agent economic environment, as well as new complications due to the endogenous policy determination.

On a conceptual level, we need to adapt the fixed-point iteration procedure used in computing RCE to account for the best-response consistency of off-equilibrium paths. As already intimated in our equilibrium definition, our procedure iterates on the best response transition function (H) and public policy function (Ψ) to reach a fixed point. Both on and off equilibrium restrictions are honored in every step of the computation.

On a practical level, we have to specify a set of moments of the wealth distribution and functional forms for (H, Ψ) to implement the general procedure proposed in Krusell and Smith (1998). They find that average capital is sufficient to approximate the infinite-dimensional wealth distribution and its law of motion and, thus, to forecast future prices. In our case, because of the public choice nature of fiscal policy higher-order statistics might matter for the evolution of the economy. This intuition is verified in our simulations. We find that the combination of average capital and the Gini coefficient of the capital distribution is sufficient to characterize the evolution of our economy. At the same time these two moments keep the dimensionality of the problem tractable.[5] We thus show that the principal idea of Krusell and Smith (1998), namely approximating the wealth distribution and its law of motion by a finite number of moments, can also be applied to politico-economic equilibrium models with uninsurable labor income risk and aggregate shocks. Notice that H has to have good predictive power not only on-equilibrium, but also for a grid of off-equilibrium proposals for G. The computed fixed

[5] We also experimented with the standard deviation of the wealth distribution, but found better R2 improvements with the Gini coefficient.

point of H then takes the following form:[6]

$$\log K' = a_0(z) + a_1(z)\log K + a_2(z)\log Gini(k) + a_3(z)\log G + a_4(z)\left(\log G\right)^2, \quad (9)$$

$$\log Gini(k') = \tilde{a}_0(z) + \tilde{a}_1(z)\log K + \tilde{a}_2(z)\log Gini(k) + \tilde{a}_3(z)\log G + \tilde{a}_4(z)\left(\log G\right)^2, \quad (10)$$

and that of Ψ:

$$\log G = b_0(z) + b_1(z)\log K + b_2(z)\log Gini(k). \quad (11)$$

Notice that the parameters of these equations depend on the (discrete) level of aggregate productivity. We solve the MPE using a fixed point iteration procedure from the parameters in (9)-(11) onto themselves. The computational algorithm is outlined in detail in Appendix C (see also Corbae et al., 2009; our procedure extends theirs by allowing for aggregate shocks). The parameterized Krusell and Smith rules and the equation-by-equation R2-statistics are available from the authors upon request.

3.5 Calibration

The model is calibrated to match features of the U.S. economy from 1960 to 2006. Annual data on government purchases correspond closely to the yearly nature of government budgeting and therefore we calibrate our model to this frequency. This choice implies three parameter selections: the depreciation rate, δ, is set to 0.1; the discount rate, β, is centered around 0.96; and we model aggregate productivity, z, as a five-state Markov chain that approximates a log-AR(1) process with an autocorrelation coefficient of 0.8145 and conditional standard deviation of 0.0165. This standard deviation is chosen to make our models approximately match the annual percentage standard deviation of GDP in the data, 1.90%. Throughout the paper for both actual data and model-simulated data we use a Hodrick-Prescott filter with smoothing parameter 100 to detrend. This paper is not about explaining output volatility from a measured exogenous shock series, as in the RBC tradition which uses fluctuations in the Solow residual to generate a large part of observed output fluctuations. Rather, this paper is concerned with shedding light on the comovement properties of government purchases and aggregate economic activity, given the right output volatility. The final standard parameter is the output elasticity of capital, $\alpha = 0.36$.

Idiosyncratic labor efficiency, ϵ, is modeled as a nine-state Markov chain that approximates a log-AR(1) process with an autocorrelation coefficient of 0.75 and conditional standard devia-

[6] It turns out that using $\left(\log G\right)^2$ improves the fit of the law of motion for capital significantly. The R2 for (9)-(10) are all above 0.999, for (9) independently of whether the Gini coefficient was used. The R2 for (11) are above 0.996, but here the introduction of the Gini coefficient matters. For instance, the R2 for (11) without the Gini coefficient is only 0.985 for the lowest aggregate productivity state.

tion of 0.18. These numbers are broadly consistent with the estimates from Guvenen and Smith (2010) who use an indirect inference approach and data on labor income, labor supply and consumption to estimate a model for the natural logarithm of labor income. We set $\tilde{l} = 0.33$.

We assume that the discount factor evolves according to a persistent three-state Markov chain pinned down by four conditions: 1) at every point in time the majority of the population (80%) has $\beta = 0.96$, and the very patient and very impatient agents have a mass of 10% each; 2) the average duration of a given discount factor is 50 years, which is meant to capture a dynastic element in this infinite horizon model; 3) agents do not jump over a state; 4) the equidistant difference between the three grid points is calibrated jointly with the borrowing constraint, $\underline{k}$, to be broadly consistent with the fraction of households with negative wealth in U.S. data and the Gini coefficient of the U.S. wealth distribution. This calibration strategy as well as its targets, 11% for the fraction of negative net wealth holders and 0.79 for the Gini coefficient, is taken from Krusell and Smith (1998) and adapted to the annual frequency. We find that given the above labor income process a small borrowing constraint of 0.01 and the following grid for β is broadly consistent with the calibration targets: $[0.94, 0.96, 0.98]$. We also find, like Krusell and Smith (1998), that other moments of the wealth distribution are matched rather well, too, such as the wealth share of the top 10 percent and the top 20 percent as well as the bottom 40 percent of wealth holders. Appendix D displays the exact specifications of the three Markov chains for z, ϵ and β.

Two parameters remain to be calibrated: θ, the weight on government purchases; and χ, the exponent of the weighting function in the political aggregation mechanism. For θ we adopt the following strategy: given the set of parameters above and a value for χ, we choose θ so that the model matches the time-averaged $\frac{G}{Y}$-ratio based on aggregate nondefense government purchases, GND, i.e. roughly 15%. We thus follow the empirical literature that views only defense government purchases as a truly exogenous stochastic process and uses this assumption to estimate the economic consequences of government spending shocks. Of course, since this paper is about endogenous government purchases we simply take the complement, i.e. nondefense government purchases. Appendix D provides the values of θ for the various models.

To calibrate χ, we must first take a stance on the source of the *de facto* political influence in the political aggregation process of the U.S. Our baseline will be a "One-Dollar-Of-Contribution-One-Vote"-interpretation of the political system, where by vote we mean the de facto influence of an agent in the political aggregation process, not her vote on election day. With such an interpretation and using the microfoundation result from the probabilistic voting literature we can map the weights in the social welfare function (8) into something observable. As a more conservative alternative we will also employ a "One-Contributor-One-Vote"-interpretation of the political system, which is essentially an extensive margin version of our baseline case: if an agent makes a strictly positive campaign contribution, then she will get one de facto vote. In the baseline case, in contrast, the votes that are allocated to the agent are proportional to her campaign contribution. That is, under the baseline case the functional form assumption, $(k_i^+)^\chi$, for the weight in the social welfare function (8) is a parametric theory of how an agent's wealth maps into campaign contributions and, therefore, de facto votes. Under the alternative interpretation, "One-Contributor-One-Vote", $(k_i^+)^\chi$ maps an agent's wealth into the probability of making a strictly positive campaign contribution.

To implement this calibration strategy for χ we draw on two different data sources. For the baseline "One-Dollar-Of-Contribution-One-Vote"-interpretation we have to use the Cooperative Congressional Election Survey (CCES) from the years 2006 and 2008. The CCES is a more detailed, but only recently developed extension of the American National Election Studies (ANES), which contains information about how much individuals donated to political campaigns. The ANES covers, biannually, the whole post-war period. It is the ANES that Rosenstone and Hansen (1993), Benabou (2000) and others have used to document political inequality in the U.S. The ANES, however, does not ask about how much an individual contributed to a campaign, only whether she did so. The ANES reports income data in five percentiles. For the "One-Contributor-One-Vote"-interpretation we use the ANES data from 1960 on, as they capture a much longer time horizon. We take the time series average of the fractions of campaign contributions or campaign contributors in each of the ANES income percentiles as the data counterpart of the time-averaged social welfare function weights from the model simulations.

There is a further complication: the ANES (but not the CCES) reports income data only in five percentiles which are not equispaced: $[0-16\%]$, $[17-33\%]$, $[34-67\%]$, $[68-95\%]$ and $[96-100\%]$. For comparability reasons we use these income percentiles and map the CCES into them. The CCES in turn reports income data in 14 classes characterized by absolute income intervals. Campaign contributions, in contrast, are reported in individual absolute dollar amounts, not as class data. When computing the ANES percentiles for the CCES data, sometimes entire income classes fall onto a percentile and the campaign contributions of these income classes have to be allocated into the neighboring groups to the left and right of a certain percentile. We consider two extreme assumptions to deal with the partially discrete nature of the data. In the baseline, conservative scenario (Scenario I) we allocate the total campaign contributions of a CCES income class that falls onto an ANES percentile half and half to the neighboring groups. This assumes essentially that within a CCES income class the campaign contributions have no systematic correlation with the underlying income level. The other extreme scenario (Scenario II) assumes the opposite, namely that the highest campaign contributions within a CCES income class must have come from the highest incomes in that income class: we order the campaign contributions in the affected income classes and attribute the total of the lower half of campaign contributions to the lower ANES income percentile and the total of the upper half of campaign contributions to the upper ANES income percentile. We do this for both CCES years, 2006 and 2008, and take the simple average as the data counterpart of the time-averaged social welfare function weights from the model simulations.

To summarize our calibration strategy: given the values for the other parameters we find χ, while recalibrating θ in the flow utility function to keep the average $\frac{G}{Y}$-ratio roughly at 15%, so as to minimize the Euclidean norm of the difference between the time-averaged campaign contribution fractions in the ANES income percentiles and the time-averaged social welfare function weights from the model simulations. The simulation numbers come from a simulation of 1,500 periods, where the first 500 periods are discarded. The simulation uses 60,000 agents and is started from an arbitrary initial wealth distribution. All simulations use the same series of aggregate shocks. Table 1 displays the results of this calibration.

Table 1: POLITICAL PARTICIPATION PER INCOME GROUP - DATA AND MODEL

Income Percentile	[0 – 16%]	[17 – 33%]	[34 – 67%]	[68 – 95%]	[96 – 100%]
"One-Dollar-Of-Contribution-One-Vote" Scenario I					
Weight shares - $\chi = 0.55$ (BL)	2.8%	6.9%	25.1%	41.8%	23.4%
Campaign contribution shares (CCES)	4.2%	7.7%	17.3%	53.1%	17.7%
"One-Dollar-Of-Contribution-One-Vote" Scenario II					
Weight shares - $\chi = 0.8$	1.4%	4.1%	17.6%	40.3%	36.5%
Campaign contribution shares (CCES)	2.9%	6.7%	16.5%	38.8%	35.1%
"One-Contributor-One-Vote"					
Weight shares - $\chi = 0.4$	4.0%	9.1%	29.4%	40.6%	16.9%
Campaign contributor shares (ANES)	5.0%	8.1%	27.5%	41.9%	17.5%

Notes: 'Income Percentile' refers to the somewhat nonstandard partition of the income distribution given by ANES. "One-Dollar-Of-Contribution-One-Vote" refers to a situation, where one dollar of campaign contributions "buys" one de facto vote in the political process. 'Scenario I' refers to a treatment of the data where the CCES campaign contributions are split equally between neighboring income percentiles. 'Scenario II' refers to a treatment of the data where the highest CCES campaign contributions in an income class are attributed to the next highest ANES income percentile if that CCES income class falls onto an ANES income percentile. "One-Contributor-One-Vote" refers to a situation, where the fact that one contributes to a political campaign "buys" this contributor one de facto vote in the political process. 'Weight shares' refers to the (time-averaged) weight shares in the social welfare function for each ANES income percentile from the model simulations. 'BL' refers to our baseline calibration $\chi = 0.55$. 'Campaign contribution shares' (from the CCES) are the average fractions of campaign contributions (from 2006 and 2008) that were made by members of a certain percentile of the income distribution. 'Campaign contributor shares' (from the ANES) are the average fractions of campaign contributors that belong to a certain percentile of the income distribution. Due to missing data problems at the recent end we use the ANES data from 1960-2000.

The way to read Table 1 is the following: the richest 5% give 17.7% of campaign contributions and represent 17.5% of campaign contributors. In contrast, the poorest 16% make up only 5% of campaign contributors and give 4.2% of all contributions. It is interesting to note that the main difference between the distribution of campaign contributors and the distribution of campaign contributions (Scenario I) lies in the difference between the lower and upper middle class, i.e. the difference between the [34−67%] and the [68−95%] bracket, whereas the difference between Scenario I and Scenario II is mainly between the upper middle class and the rich. It should also be intuitively clear that calibrating to the distribution by income of the extensive margin of campaign contributions, the campaign contributors (the last panel in Table 1), requires a lower pro-wealth bias ($\chi = 0.4$) than calibrating to the distribution by income of the intensive margin of campaign contributions, which takes into account the size of the contributions made by the different income groups – $\chi = 0.55$ in the conservative scenario, which can be found in the first panel of Table 1, and $\chi = 0.8$ in the other extreme scenario, which can be found in the second panel of Table 1.[7]

We view this evidence as suggestive that there is enough pro-wealth bias in the U.S. political system to be quantitatively relevant. In Section 4 we nevertheless report results for a whole range of χ. That section will show that the various calibration scenarios entertained here include the range of pro-wealth bias that is consistent with the amount of decoupling between government purchases and aggregate economic activity observed in the data, both for a social welfare function / probabilistic voting as well as a majority voting political aggregation mechanism.

4 Results

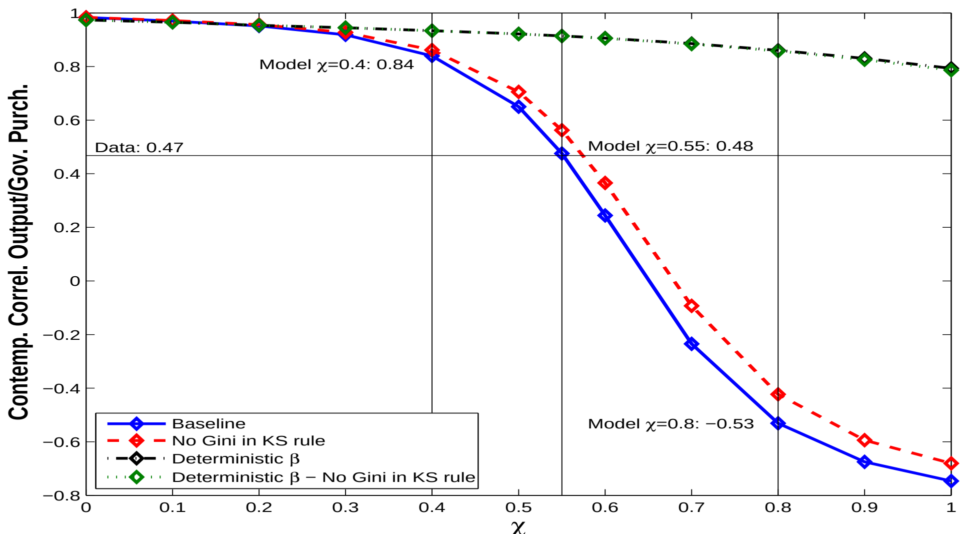
4.1 The Baseline Result

Decoupling between government purchases and the overall macroeconomic activity increases with the degree of wealth-bias in the political system, parameterized by the exponent in the political weight function, $(k_i^+)^\chi$. Starting from the case with no wealth bias in the social welfare function ($\chi = 0$) – every agent in the economy has the same political weight – we increase χ and plot in Figure 1 the contemporaneous correlation coefficient between output and government purchases. We do this for the baseline computation/calibration and three variants: first, we leave out the Gini coefficient in the Krusell-Smith rules for capital and government purchases (“No Gini in KS rule” - dashed line), and recompute the equilibrium; secondly, we do not tar-

[7] Neither the ANES nor the CCES have data on wealth, so we have to use political inequality by income percentile as a proxy for a pro-wealth bias in the political system.

get a realistic Gini coefficient in the simulated wealth distribution and set the discount factor deterministically to its median value of 0.96, but leave the higher moment in the Krusell-Smith rules (dashed-dotted line); thirdly, we combine both changes (dotted line).[8]

Figure 1: Correlation Between Y and G as a Function of Wealth Bias in the Baseline Calibration



Notes: In the simulations all variables are logged and filtered with a Hodrick-Prescott filter with smoothing parameter 100. 'No Gini in KS rule' refers to the equilibrium, when we leave out the Gini coefficient in the Krusell-Smith rules for capital and government purchases. 'Deterministic β' refers to the case of a deterministic discount factor set to its median value 0.96, but with the Gini coefficient in the Krusell-Smith rules for capital and government purchases. 'Deterministic β - No Gini in KS rule' refers to the equilibrium with deterministic β and no Gini coefficient in the Krusell-Smith rules. 'Model $\chi = 0.4$' is the model that matches best the political influence data under a "One-Contributor-One-Vote"-interpretation. 'Model $\chi = 0.55$' is the model that matches best the political influence data under a "One-Dollar-Of-Contribution-One-Vote"-interpretation, Scenario I. 'Model $\chi = 0.8$' is the model that matches best the political influence data under a "One-Dollar-Of-Contribution-One-Vote"-interpretation, Scenario II.

The comovement between output and government purchases declines with the degree of wealth bias in the baseline case. At a value of $\chi = 0.55$ (the "One-dollar-one-vote"-interpretation, Scenario I) the model matches the observed contemporaneous comovement between government purchases and output almost exactly. Comparing the baseline simulation with the Gini coefficient in the Krusell-Smith rules and the recomputed equilibrium without any higher moments reveals a quantitatively realistic example where there is at least a small deviation from approximate aggregation in the sense of Krusell and Smith (1998). Using a measure of wealth inequality in the equilibrium law of motion changes actual equilibrium dynamics slightly for higher political bias parameters - the solid and the dashed line deviate from each other for

[8]The resulting Gini coefficient is 0.46 and the fraction of negative wealth holders is 2.8% across models, whether we use the Gini coefficient or not in the Krusell-Smith rules.

higher χ. When agents do not take into account the dynamics of wealth inequality in their forecasting rules, the decoupling effect that a highly unequal wealth distribution in concert with a wealth-biased political system brings about is slightly mitigated. The contemporaneous correlation between output and government purchases increases from 0.48 to 0.56 for the case of $\chi = 0.55$, when the Gini coefficient is left out. Figure 1 also shows that with a counterfactual wealth distribution and only mild economic inequality the effect of wealth bias on aggregate government purchases dynamics is nearly eliminated, whether we include higher moments in the forecasting rules or not. Finally, in the case of complete political equality, $\chi = 0$, the extent of economic heterogeneity is irrelevant for aggregate comovement: all lines in Figure 1 have their origin in the same point at $\chi = 0$.

Table 2: BASELINE SIMULATION RESULTS - $\chi = 0.55$

Moment	Correl. w. Y	Correl. w. Y-Lag	Correl. w. C	Autocorrel. 1st-order	Vol.
Y	1.00 (1.00)	0.51 (0.54)	0.93 (0.87)	0.51 (0.54)	1.88 (1.90)
C	0.93 (0.87)	0.69 (0.41)	1.00 (1.00)	0.65 (0.62)	1.08 (1.67)
I	0.97 (0.84)	0.34 (0.21)	0.81 (0.69)	0.45 (0.42)	5.48 (7.84)
GND	0.48 (0.47)	0.74 (0.58)	0.75 (0.49)	0.85 (0.74)	0.58 (1.87)

Notes: For an explanation of the model simulations see notes to Figure 1. Data moments are in brackets. 'Vol.' is the percentage standard deviation of a variable. Y denotes GDP, C private consumption expenditures and I private gross fixed investment. GND denotes government nondefense consumption and gross investment expenditures. All variables are annual, they range from 1960-2006. They are deflated by their corresponding deflators, logged and filtered with a Hodrick-Prescott filter with smoothing parameter 100.

Table 2 shows that at $\chi = 0.55$ not only the mildly positive contemporaneous correlation between government purchases and output is matched (data moments are in brackets), but also the contemporaneous correlation between government purchases and private consumption is relatively low, and the same holds true for the correlation between government purchases and lagged output. Our model is also consistent with government purchases being the most persistent component of aggregate demand.[9] As it is in the data, government purchases are also more persistent than aggregate output. Table 2 also shows that the model only generates roughly one third of the volatility of government purchases in the data, which is essentially as volatile as output. Below we will show that the model features an amplification-propagation

[9] The model overshoots, but GND is one of the least persistent government purchases aggregates. Had we calibrated to the $\frac{G}{Y}$-ratio based on G (22.5%) we would have matched the high persistence almost perfectly. Tables 11 and 12 in Appendix E display comovement and persistence for a variety of disaggregate measures of government purchases. They show that the GND-aggregate is typical: most disaggregations of G, both according to administrative units and according to a functional disaggregation, have low but positive contemporaneous correlations with output and private consumption, a higher dynamic correlation, and a persistence that exceeds that of aggregate output.

trade-off along χ, i.e., when the model-generated first-order autocorrelation for G is high, as it is in the case of $\chi = 0.55$, volatility of G is low.

We use a Hodrick-Prescott filter with smoothing parameter 100 to detrend both the actual data and the simulated data from the model. The correlation pattern we find is not sensitive to this choice. Persistence of government purchases is much lower with a smoothing parameter of 6.25. However, we do not want to limit our analysis to business cycle fluctuations, narrowly defined. Our mechanism – time series movements of political inequality – is likely a process with important medium-run components. That government purchases in the data exhibit important and very persistent medium-run fluctuations is consistent with this idea. Hence the choice of a "stiffer" trend.

4.2 Explanation

How do economic and political heterogeneity interact to generate the decoupling result?

Table 3: SIMULATION RESULTS - THE ROLE OF ECONOMIC AND POLITICAL HETEROGENEITY

Moment	Baseline ($\chi = 0.55$)	$\chi = 0.4$	$\chi = 0.8$	$\chi = 0$	Determ. β $\chi = 0.55$	Rep. Agent	Data
G							
Correl. w. Y	0.48	0.84	-0.53	0.98	0.91	0.96	0.47
Correl. w. Y-Lag	0.74	0.73	0.22	0.60	0.70	0.66	0.58
Correl. w. C	0.75	0.98	-0.22	0.98	1.00	0.98	0.49
Autocorrel. 1st-order	0.85	0.72	0.55	0.57	0.67	0.62	0.74
Vol.	0.58	0.75	0.92	1.45	0.83	1.08	1.87

Notes: see notes to Table 2. 'Baseline' refers to the heterogeneous agent model with wealth bias parameter $\chi = 0.55$, it fits the cross-sectional data on political influence best under a "One-Dollar-Of-Contribution-One-Vote"-interpretation (Scenario I) of the political system. $\chi = 0.4$ and $\chi = 0.8$ refer to models that fit the cross-sectional data on political influence best under a "One-Contributor-One-Vote"-interpretation and a "One-Dollar-Of-Contribution-One-Vote"-interpretation (Scenario II), respectively. $\chi = 0$ is a model without any wealth bias, i.e. without any political inequality, in the social welfare function. 'Determ. β' refers to the case of a deterministic discount factor set to its median value 0.96, but with the Gini coefficient in the Krusell-Smith rules for capital and government purchases. We essentially keep the same parameter for wealth bias as in the baseline case, i.e. $\chi = 0.55$, but reduce the extent of wealth inequality. 'Rep. Agent' refers to a simulation, where we abstract from any heterogeneity, i.e. ϵ and β are held constant.

Table 3 compares the dynamics of government purchases from the baseline calibration with those from alternative calibrations of the pro-wealth bias in the political system (second vertical panel) and with those from models that do not generate decoupling at all, i.e. models that either lack political inequality, sufficient economic inequality or both (third vertical panel).

The case with $\chi = 0.4$, the "One-Contributor-One-Vote"-interpretation, leads to less decoupling with output (no decoupling with private consumption), less propagation and more amplification of government purchases than the baseline case. The case with $\chi = 0.8$, "One-Dollar-Of-Contribution-One-Vote"-interpretation (Scenario II), features more decoupling with output and private consumption, less propagation and more amplification of government purchases than the baseline case.

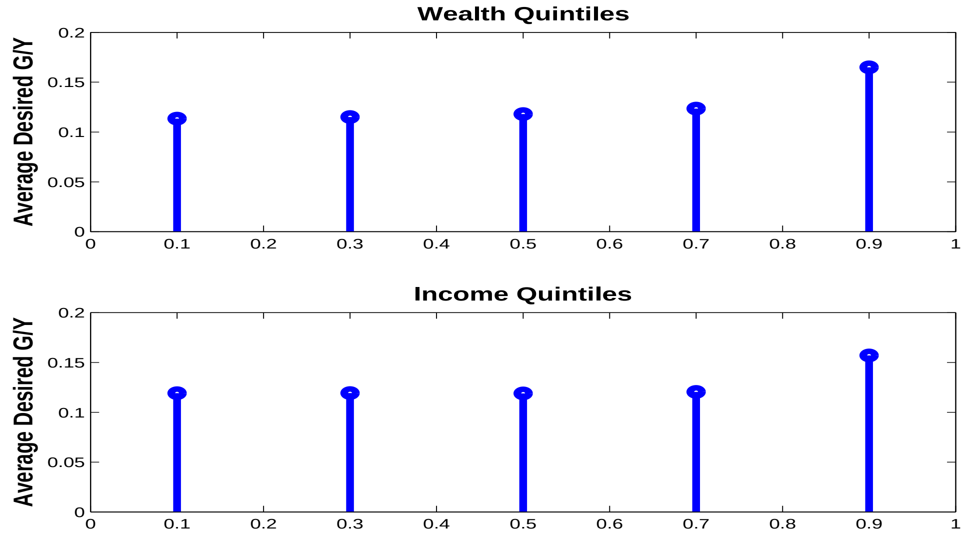
Models with either a lack of realistic economic inequality (independently of the extent of wealth bias) – the 'Determ. β'-column in the table – or with a lack of political inequality (independently of whether they feature realistic economic inequality) – the $\chi = 0$-column in the table – behave in the aggregate essentially like representative agent versions of the model without any heterogeneity and inequality – the 'Rep. Agent'-column in the table. This finding also holds for other moments and other aggregate variables not shown in Table 3.

We thus provide another case of irrelevance of wealth heterogeneity for aggregate dynamics – government purchases dynamics, to be precise –, extending the finding of Krusell and Smith (1998) to a class of models with endogenous public policy. However, once both economic and political heterogeneity are combined, they matter for the aggregate dynamics of government purchases. Because of this, government purchases dynamics provide important restrictions on a difficult-to-measure structural parameter, namely the wealth bias in the political system, which suggests that future empirical research might benefit from taking these dynamics into account. Only the baseline model with both economic and political inequality generates the right level of positive comovement of government purchases and aggregate economic activity. With a fixed discount factor, economic inequality is too small to matter. While the 'Determ. β'-case with lower wealth inequality also has a negative relationship between wealth bias and the comovement of GDP with government purchases, the gradient is small. With $\chi = 0$, economic inequality does not get translated into political inequality.

Table 3 also shows that the specification with both types of inequality has other important moments closer to the data, in particular the correlation of government purchases with private consumption and the persistence of government purchases. In the baseline specification, G is the most persistent component of aggregate demand. This is consistent with the fact that government purchases dynamics now depend directly on a persistent object, the wealth distribution. This comes, however, at the price of little amplification of aggregate shocks into government purchases fluctuations. There is a propagation-amplification trade-off for government purchases with respect to the pro-wealth bias parameter, χ.

Next, we show that the decoupling we find in the baseline simulation is caused by a shift of agents' political power over time. Figure 2 shows that the average desired G increases with wealth and income quintiles. Figure 3 shows that the relative political influence of the poor

Figure 2: Desired Government Purchases per Wealth/Income Quintile



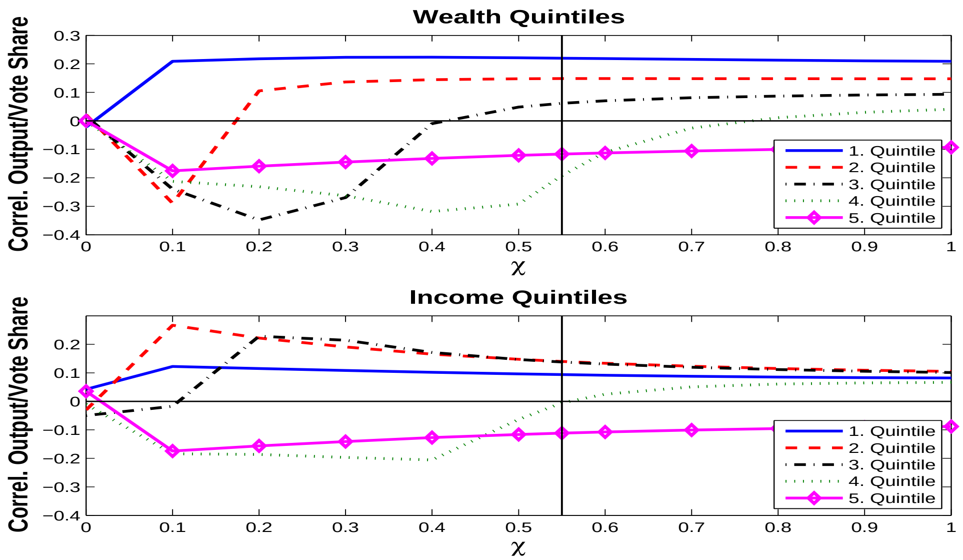
Notes: We use the simulation with $\chi = 0.55$ and $\theta = 0.815$. We compute for each of the 60,000 agents at every point in the simulation their desired level of government purchases, using their indirect net present value utility function over government purchases, J_i. With this we can compute the conditional averages of desired government purchases per wealth or income quintile. We normalize these conditional means by the level of aggregate output at that point in time. The figure plots the time averages of these numbers.

increases in good economic times, whereas the rich lose some of their weight in the political decision process. In good economic times the poor have slightly more say and they want on average lower government purchases, which mitigates the expansion of G. Figure 3 also shows that only very mild positive comovement of the political influence of the poor and mild negative comovement of the political influence of the rich are required to generate the extent of decoupling observed in aggregate government purchases. In the ANES data, the lowest income percentile's voting share has a correlation with GDP of 0.14 and 0.24 with GND; for the highest income percentile these numbers are -0.05 and -0.21, respectively. The lowest income percentile's campaign contributions has a correlation with GDP of 0.08 and 0.22 with GND; for the highest income percentile these numbers are -0.22 and -0.11, respectively.[10]

Why is political inequality countercyclical in this model? Because wealth inequality is. The wealth share of the lowest wealth quintile in the baseline model ($\chi = 0.55$) is positively correlated with GDP (0.19) and the wealth share of the highest wealth quintile is negatively correlated with GDP (-0.08). The Gini coefficient of wealth in the baseline model is negatively correlated with GDP (-0.15). Intuitively, countercyclical wealth inequality is the result of two forces: first,

[10]Since the CCES data are not long enough, we cannot report these business cycle statistics for the distribution of the intensive margin of campaign contributions.

Figure 3: Comovement of Political Influence per Wealth/Income Quintile as a Function of the Wealth Bias



Notes: For every simulation, as χ varies, we compute the weight shares in the social welfare function for each wealth or income quintile. We then compute for each wealth or income quintile's weight share the contemporaneous correlation coefficients with model simulated output, which is logged and filtered with a Hodrick-Prescott filter with smoothing parameter 100.

the right-hand side of the households' budget constraint, (4), shows that an increase in aggregate productivity z increases the fraction of income, i.e. the flow component (second summand), in the total available resources compared to the stock component of wealth carried over (first summand). This means that as long as the flow component (income) is more equally distributed than the stock component (wealth), as is the case both in the data and in the model, total available resources of the households will be more equally distributed in a boom (see Appendix F for a more technical argument). Suppose, secondly, that the optimal saving policy is a linear function of total available resources, which is approximately the case in these Aiyagari-type models, especially for rich households (see Krusell and Smith, 1998), then greater equality in total available resources today will translate into greater equality of wealth tomorrow. With persistent productivity shocks this means countercyclical wealth inequality.

The next Table 4 explores whether the specific way we model wealth inequality, the stochastic discount factor approach, is essential to generate the decoupling result. To this end, we follow the method by Domeij and Heathcote (2004, pp 548-550) and use an extremely skewed, three-state labor productivity process to match important features of the labor income process and the wealth distribution, with calibration targets tailored to our application.[11] Just as in the

[11] The basic idea to use the richness of free parameters in a Markov chain to calibrate the labor income process and the wealth distribution jointly was introduced by Castaneda et al. (2003) into the literature, but the specific application in Domeij and Heathcote (2004) fits our purposes better. Given the assumptions in Domeij and Heath-

Table 4: SIMULATION RESULTS - THE ROLE OF THE WEALTH DISTRIBUTION

Moment	Baseline ($\chi = 0.55$)	Ext. Income ($\chi = 0.4$) Top 10 Wealth Share	Ext. Income ($\chi = 0.4$) Top 10 Wealth Share Too Low	Ext. Income ($\chi = 0.4$) No Borrowing	Data
G					
Correl. w. Y	0.48	0.66	0.93	0.95	0.47
Correl. w. Y-Lag	0.74	0.48	0.39	0.66	0.58
Correl. w. C	0.75	0.62	0.83	0.98	0.49
Autocorrel. 1st-order	0.85	0.65	0.40	0.62	0.74
Vol.	0.58	0.97	0.95	0.78	1.87

Notes: see notes to Table 2. 'Baseline' refers to the heterogeneous agent model with wealth bias parameter χ = 0.55, it fits the cross-sectional data on political influence best under a "One-Dollar-Of-Contribution-One-Vote"-interpretation (Scenario I) of the political system. 'Ext. Income Top 10 Wealth Share' refers to a calibration, where we fix β at 0.96, and generate the wealth inequality in the data by using an extremely skewed income process. We recalibrate χ to 0.4 under a "One-Dollar-Of-Contribution-One-Vote"-interpretation (Scenario I). This scenario matches not only the fraction of negative wealth and the Gini coefficient of wealth in the data, but also the wealth share of the top 10 percent of the wealth distribution, 0.64. 'Ext. Income Top 10 Wealth Share Too Low' is the same as the preceding calibration, except that we now match a counterfactual wealth share of the top 10 percent of the wealth distribution that is 0.1 lower than in the data. 'Ext. Income No Borrowing' is the same as 'Ext. Income Top 10 Wealth Share', except that we do not allow for borrowing. This scenario is roughly consistent with the Gini coefficient and the wealth share of the top 10 percent of the wealth distribution in the data, but no longer with the fraction of negative wealth holders.

baseline case, we target both the persistence (0.75) and standard deviation (0.18) of the innovation of the underlying AR(1) labor income process, as well as the Gini coefficient (0.79). We also target an additional statistic of the wealth distribution, the wealth share of the top 10 percent wealth-richest households (0.64 as reported in Krusell and Smith, 1998). It turns out that matching the wealth distribution at the upper end is crucial for the decoupling result. These targets together with the fraction of negative wealth holders (11%) are approximated in the general equilibrium computation by calibrating the labor income process and the borrowing constraint jointly, with β fixed at 0.96. Finally, we recalibrated the wealth bias parameter to $\chi = 0.4$ under the "One-Dollar-Of-Contribution-One-Vote"-interpretation (Scenario I). Appendix D summarizes the exact parameterizations of the calibrated Markov chains.

Column 2 of Table 4 shows that generating a quantitatively realistic wealth distribution with the Domeij and Heathcote (2004) approach leads to somewhat less decoupling of government purchases compared to our baseline where the Krusell and Smith (1998) approach was used.

cote (2004), there remain essentially four parameters of the three-state Markov chain to be calibrated.

The difference between the two cases is entirely due to the recalibration, as the cross-sectional evidence on political inequality under a “One-Dollar-Of-Contribution-One-Vote”-interpretation (Scenario I) can be best replicated with a slightly lower χ-parameter.[12] Column 3 and 4 of the same table show that what matters for these results is the influence of the upper part of the wealth distribution, specifically the wealth share of the top 10 percent of the wealth distribution (column 3),[13] as well as the influence of the fraction of borrowers (column 4). In column 3, we counterfactually calibrate to a top 10 wealth share of 0.54, but recalibrate the model to remain consistent with the Gini coefficient of wealth and the fraction of borrowers in the data. In column 4, we shut down all borrowing exogenously, but recalibrate the model to remain consistent with the Gini coefficient of wealth and the top 10 share of the wealth distribution. In both cases, the models feature essentially no decoupling. The upshot of this investigation is that at least qualitatively the decoupling effect is robust to using a different way of generating an unequal wealth distribution.

4.3 Wealth-weighted Majority Voting

In this section, we study a case where we use (wealth-weighted) majority voting as the political constitution with vote allocation for type (k,ϵ,β) equal to $(k^+)^\chi$. The baseline model featured a (wealth-weighted) social welfare function as the political aggregation mechanism, which can be shown to be equivalent to a (wealth-weighted) probabilistic voting mechanism. As in a standard majority voting environment, the policy with fifty percent of the votes is the winner if the preference of every household is single-peaked.[14]

Specifically, let $\widehat{G}\left(k,\epsilon,\beta,\Gamma,z;\Psi,H\right) = \arg\max_G J\left(k,\epsilon,\beta,\Gamma,z,G;\Psi,H\right)$ be the preferred policy of economic type $\left(k,\epsilon,\beta\right)$. Then the Condorcet winner in a weighted majority voting system is the preferred policy of type $\left(k^*,\epsilon^*,\beta^*\right)$ such that

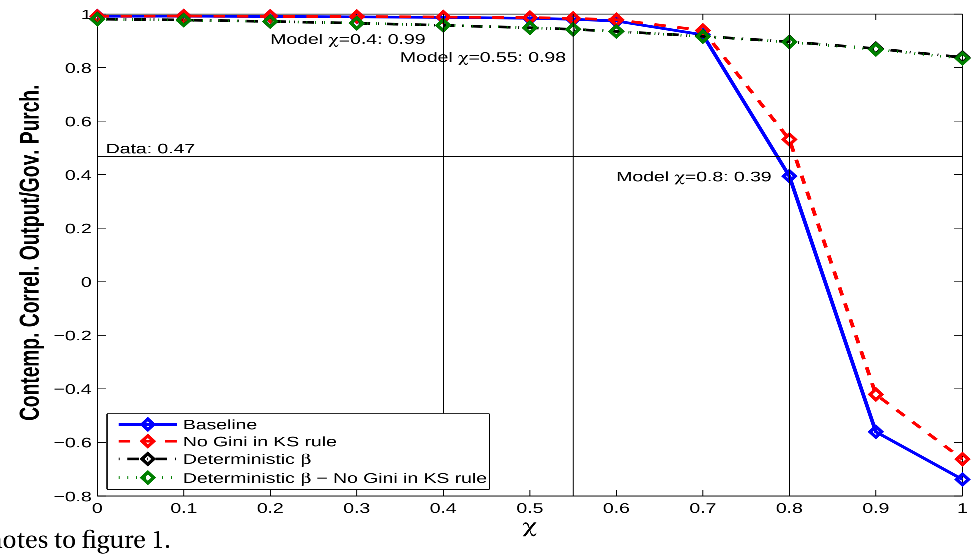
$$\frac{\int_{\{i:\widehat{G}(k_i,\epsilon_i,\beta_i)<\widehat{G}(k^*,\epsilon^*,\beta^*)\}}(k_i^+)^\chi di}{\int(k_i^+)^\chi di} = \frac{1}{2}.$$

[12] In fact, in the case of $\chi = 0.55$ the aggregate statistics for government purchases would look much like those in the baseline case where the Krusell and Smith (1998) approach was used.

[13] In the baseline Krusell and Smith calibration we followed the individual capital holdings of a household with median productivity, ϵ, and median patience, β, whose individually optimal level of government purchases was closest to the actually chosen aggregate level of government purchases. At $\chi = 0.55$ (with sizeable decoupling) the individual capital holdings of said household is just outside the upper 10 percent of the wealth distribution. In contrast, at $\chi = 0.40$ (with less decoupling), the “One-Contributor-One-Vote”-interpretation, said household is just inside the upper 20 percent of the wealth distribution.

[14] In the voting literature, there are different sufficient conditions to guarantee the existence of the voting equilibrium, e.g., single-peakedness, intermediate preference and single crossing (or equivalently order restriction) conditions (see Persson and Tabellini, 2000, for an overview). Although there are general existence results for the complete-market neoclassical growth model with policy commitment(see Bassetto and Benhabib, 2006), we are not aware of similar results applicable to our environment. Nevertheless, we verify numerically that single-peakedness is satisfied in our simulations.

Figure 4: Correlation Between Y and G as a Function of Wealth Bias - Wealth-weighted Majority Voting



Notes: see notes to figure 1.

Figure 4 repeats Figure 1 for the case of majority voting, where again higher political bias yields decoupling of government purchases and the aggregate economy. For the majority voting case the slope of the decline of the contemporaneous correlation coefficient between output and government purchases in χ is much flatter initially and much steeper for higher values of χ, compared to the probabilistic voting / social welfare function case. As in the baseline case, the "One-Contributor-One-Vote"-interpretation leads to $\chi = 0.4$, the "One-Dollar-Of-Contribution-One-Vote"-interpretation (Scenario I) leads to $\chi = 0.55$ and the "One-Dollar-Of-Contribution-One-Vote"-interpretation (Scenario II) leads to $\chi = 0.8$, which is also the level of pro-wealth bias, where the contemporaneous correlation between GDP and government purchases is close to the data. This shows that a level of decoupling that is consistent with aggregate time series data on government purchases can be generated under a constitution that is different from probabilistic voting, if one is willing to assume a level of political inequality in the upper range of what is consistent with the cross-sectional evidence on the pro-wealth bias of political influence.

5 Final Remarks

This paper provides a quantitative theory of the implications of economic and political inequality for the aggregate dynamics and, specifically, the comovement of government purchases with overall economic activity. We show that with both types of inequality present in a quantitatively realistic way, the interaction of government purchases being a normal good and negatively comoving political inequality can alter the time series behavior of government purchases compared to models where either type of inequality is missing. An extreme case of the latter is a representative agent environment. We argue that in standard models the positive comovement of government purchases with overall economic activity may be dampened in the presence of economic and economically-based political inequality. This finding does not depend on the specifics of the political system. In contrast, standard neoclassical representative agent models of endogenous public policy feature too much comovement of government purchases. We thus provide an example of a quantitatively realistic model, where heterogeneity matters for aggregate dynamics, specifically the dynamics of government purchases.

This does, of course, not mean that ours is the only conceivable mechanism that could explain the mild positive comovement of government purchases. Given the fact that most quantitatively sizeable disaggregates of government purchases behave similarly to the aggregate government purchases (see Table 12 in Appendix E), it is unlikely, however, that the stylized data facts that we emphasize are merely the result of a composition effect, where some disaggregates are used to conduct countercyclical fiscal policy and others behave just as private consumption and are therefore very procyclical, and thus the aggregate ends up displaying mild procyclicality. Another explanation could be that the U.S. government has a much better access to capital markets than the U.S. private sector and therefore smoothes its consumption better. We notice that indeed the federal government's purchases are much less procyclical compared to the purchases of states and municipalities (see Table 11 in Appendix E), which may be explained by the relative better access of the federal government to borrowing compared to the states' and local authorities' access. However, better access to borrowing does not necessarily mean lower correlation with aggregate economic activity, when the shocks that are smoothed out are mainly aggregate shocks that cannot be insured against by, say, the U.S. government borrowing from the international capital markets. We leave it up to future research to determine quantitatively, whether the U.S. business cycle is sufficiently decoupled from the world business cycle and thus sufficiently insurable by the U.S. government in the international capital markets to explain the mildly positive comovement of government purchases with U.S. economic activity. The third explanation would be a second shock, which is at least partially uncorrelated with the shocks that drive aggregate economic activity, for example in the political system itself: we have pursued this idea in companion work in Bachmann and Bai (2011).

In any event, standard economic environments with one aggregate shock would have to feature enough nonlinearities to generate the observed mild positive comovement of government purchases with economic activity. This paper shows that in a standard economic environment quantitatively realistic decoupling could be the consequence of quantitatively realistic economic and political inequality.

References

[1] Alesina, A., Campante, F., Tabellini, G., 2008. Why is fiscal policy often procyclical, Journal of the European Economic Association 6, 1006–1036.

[2] Aiyagari, S.R., 1994. Uninsured idiosyncratic risk and aggregate saving, Quarterly Journal of Economics 109, 659–684.

[3] Azzimonti, M., 2011. Barriers to investment in polarized societies, American Economic Review 101, 2182–2204.

[4] Azzimonti, M., Battaglini, M., Coate, S., 2010. Analyzing the case for a balanced budget amendment to the U.S. constitution, mimeo.

[5] Bachmann, R., Bai, J.H., 2011. Public consumption over the business cycle, NBER-WP 17230.

[6] Bai, J.H., Lagunoff, R., 2011a. On the faustian dynamics of policy and political power, Review of Economic Studies 78, 17–48.

[7] Bai, J.H., Lagunoff, R., 2011b. Revealed political power, mimeo.

[8] Barseghyan, L., Battaglini, M., Coate, S., 2010. Fiscal policy over the real business cycle: a positive theory, mimeo.

[9] Bartels, L., 2008. Unequal Democracy: The Political Economy of the New Gilded Age. Princeton University Press.

[10] Bassetto, M., Benhabib, J., 2006. Redistribution, taxes, and the median voter, Review of Economic Dynamics 9, 211–223.

[11] Battaglini, M. Coate, S., 2008. A dynamic theory of public spending, taxation and debt, American Economic Review 98, 201–236.

[12] Benabou, R., 2000. Unequal societies: income distribution and the social contract, American Economic Review 90, 96–129.

[13] Boustan, L., Ferreira, F., Winkler, H., Zolt, E., 2010. Income inequality and local government in the United States, 1970-2000, NBER-WP 16299.

[14] Campante, F., 2011. Redistribution in a model of voting and campaign contributions, Journal of Public Economics 95, 646–656.

[15] Castaneda, A., Diaz-Gimenenz, J., Rios-Rull, V., 2003. Accounting for the U.S. Earnings and Wealth Inequality, Journal of Political Economy 111, 818–856.

[16] Corbae, D., D'Erasmo, P., Kuruscu, B., 2009. Politico-economic consequences of rising wage inequality, Journal of Monetary Economics 56, 43–61.

[17] Debortoli, D., Nunes, R., 2010. Fiscal policy under loose commitment, Journal of Economic Theory 145, 1005–1032.

[18] Diaz-Gimenez, J., Quadrini, V., Rios-Rull, V., 1997. Dimensions of inequality: facts on the U.S. distributions of earnings, income, and wealth, Federal Reserve Bank of Minneapolis Quarterly Review 21, 3–21.

[19] Domeij, D., Heathcote, J., 2004. On the Distributional Effects of Reducing Capital Taxes, International Economic Review 45, 523–554.

[20] Gomes, F., Michaelides, A., Polkovnichenko, V., 2008. Fiscal policy in an incomplete markets economy, mimeo.

[21] Guvenen, F., Smith, A., 2010. Inferring labor income risk from economic choices: an indirect inference approach, NBER-WP 16327.

[22] Hassler, J., Krusell, P. Storesletten, K., Zilibotti, F., 2005. The dynamics of government, Journal of Monetary Economics 52, 1331–1358.

[23] Hassler, J., Mora, J., Storesletten, K., Zilibotti, F., 2003. The survival of the welfare state, American Economic Review 93, 87–112.

[24] Heathcote, J., 2005. Fiscal policy with heterogeneous agents and incomplete markets, Review of Economic Studies 72, 161–188.

[25] Heathcote, J., Storesletten, K., Violante, G., 2010. Unequal we stand: an empirical analysis of economic inequality in the United States, 1967-2006, Review of Economic Dynamics 13, 15–51.

[26] Huggett, M., 1993. The risk-free rate in heterogeneous-agent incomplete-insurance economies, Journal of Economic Dynamics and Control 17, 953–969.

[27] Ilzetzki, E., 2011. Rent-seeking distortions and fiscal procyclicality, Journal of Development Economics 96, 30–46.

[28] Ilzetzki, E., Vegh, C., 2008. Procyclical fiscal policy in developing countries: truth or fiction?, NBER-WP 14191.

[29] Klein, P., Krusell, P., Rios-Rull, V., 2008. Time-consistent public policy, Review of Economic Studies 75, 789–808.

[30] Krusell, P., Quadrini, V., Rios-Rull, V., 1997. Politico-economic equilibrium and economic growth, Journal of Economic Dynamics and Control 21, 243–272.

[31] Krusell, P., Rios-Rull, V., 1999. On the size of U.S. government: political economy in the neoclassical growth model, American Economic Review 98, 1156–1181.

[32] Krusell, P., Smith, A., 1997. Income and wealth heterogeneity, portfolio choice and equilibrium asset returns, Macroeconomic Dynamics 1, 387–422.

[33] Krusell, P., Smith, A., 1998. Income and Wealth Heterogeneity in the Macroeconomy, Journal of Political Economy 106, 867–896.

[34] Lindbeck, A., Weibull, J., 1987. Balanced-budget redistribution as political equilibrium, Public Choice 52, 273–297.

[35] Persson, T., Tabellini, G., 2000. Political Economics: Explaining Economic Policy. MIT Press.

[36] Rosenstone, S., Hansen, J., 1993. Mobilization, Participation and Democracy in America. Macmillan, New York.

[37] Song, Z., 2012. Persistent ideology and the determination of public policy over time, International Economic Review 53, 175–202.

[38] Song, Z., Storesletten, K., Zilibotti, F., 2011. Rotten parents and disciplined children: a politico-economic theory of public expenditure and debt, mimeo.

[39] Tauchen, G., 1986. Finite state Markov-chain approximations to univariate and vector autoregressions, Economics Letters 20, 177–181.

A Proof of the Proposition - Appendix

Substitute the budget constraint (4) for private consumption in (1), use the CRRA assumption, and, suppressing the arguments in $p(k,\epsilon;K,L)$, take the derivative with respect to G:

$$FOC(G,z,k) \equiv -\theta p u_1'\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big) + (1-\theta)u_2'(G) = 0.$$

By the implicit function theorem and the strict concavity of the objective function, we need to show that $\frac{\partial FOC(G,z,k)}{\partial z} > 0$ and $\frac{\partial FOC(G,z,k)}{\partial k} > 0$ to prove the result.

$$\frac{\partial FOC(G,z,k)}{\partial z} = -\theta p^2 u_1'\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big)K^{\alpha}L^{1-\alpha} > 0$$

by strict concavity of u_1.

To prove $\frac{\partial FOC(G,z,k)}{\partial k} > 0$, we specialize $FOC(G,z,k)$ to

$$FOC(G,z,k) \equiv -\theta p\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big)^{\gamma-1} + (1-\theta)u_2'(G) = 0.$$

From this it is obvious that G is independent of z if $\gamma = 1$. Furthermore:

$$\begin{aligned}
\frac{\partial FOC(G,z,k)}{\partial k} &= -\theta p(\gamma-1)\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big)^{\gamma-2}\Big((1-\delta) + \frac{\partial p}{\partial k}(zK^{\alpha}L^{1-\alpha} - G)\Big) - \\
&\theta\frac{\partial p}{\partial k}\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big)^{\gamma-1} = -\theta\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big)^{\gamma-2} * \\
&\Big[p(\gamma-1)(1-\delta) + \frac{\partial p}{\partial k}\Big(p(\gamma-1)(zK^{\alpha}L^{1-\alpha} - G) + (1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\Big)\Big] = \\
&-\theta\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big)^{\gamma-2}\Big[p(\gamma-1)(1-\delta) + \frac{\partial p}{\partial k}\Big(p\gamma(zK^{\alpha}L^{1-\alpha} - G) + (1-\delta)k\Big)\Big] = \\
&-\theta\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big)^{\gamma-2}\Big[(p - \frac{\partial p}{\partial k}k)(\gamma-1)(1-\delta) + \frac{\partial p}{\partial k}\gamma c\Big] = \\
&-\theta\big((1-\delta)k + pzK^{\alpha}L^{1-\alpha} - pG\big)^{\gamma-2}\Big[(1-\alpha)\frac{\tilde{l}\epsilon}{L}(\gamma-1)(1-\delta) + \frac{\partial p}{\partial k}\gamma c\Big].
\end{aligned}$$

The first factor of the last expression is negative. The second factor is negative for $\gamma \le 0$ and $\delta \le 1$ with one inequality holding strict. Recall that $\frac{\partial p}{\partial k} > 0$. With linear utility in consumption, $\gamma = 1$, $\frac{\partial FOC(G,z,k)}{\partial k} < 0$.

B An Alternative Interpretation of the Model - Appendix

Recall that in the decoupling mechanism described in the main text government purchases are a normal good both over time and across agents. This may be at odds with a view of the world, where permanently more productive households are also those agents that put a lower preference weight on government purchases. The following discussion shows that such a modification would not necessarily alter our analysis. Our argument is based on the fact that the baseline setup in Section 3.1-3.3 can be alternatively interpreted as an extended model with *ex ante* heterogeneity in average individual labor productivity levels, preferences for government purchases and borrowing constraints. The crucial assumption here is that these differences are permanent; or rather that they change at a very low frequency.

To be specific, we now consider an extension to our benchmark model in Section 3.1- 3.3 with *ex ante* heterogeneous permanent household types, $j \in \mathbb{J} = \{L, H\}$:[15] before time 0 a fraction ϕ^j of households is assigned to a type j, where $\phi^j \in (0,1)$ and $\phi^L + \phi^H = 1$. Type j determines both the permanent component of labor efficiency e^j and preferences towards government purchases θ^j, with $e^L < 1 < e^H$, $\theta^L < \theta^H < 1$. We impose, given our constant returns to scale assumption, an innocuous normalization, $\sum_j \phi^j e^j = 1$, to keep the size of the benchmark economy and the size of the economy with ex ante heterogeneity the same. $j = L$ is associated with a lower permanent individual labor efficiency and a "big government"-ideology with respect to government purchases. From time 0 onwards, each household faces the same stochastic environment, (ϵ, β, z), as in our benchmark model.

We assume that a type j household with ϵ_t in period t has a multiplicative effective labor efficiency $e^j \epsilon_t$, and faces a borrowing constraint proportional to her e^j, i.e. $e^j \underline{k}$. To ease the comparison with the benchmark economy with no ex ante heterogeneity, we use lower case variables (c, k, k') to represent their counterparts normalized by e^j, e.g. c units of normalized consumption for type j implies a household consumption of $\tilde{c}^j = e^j c$. Defined in terms of normalized variables, the household flow budget equation and borrowing constraint for type j stay the same as those in Section 3.1, i.e., defined by equation (7) and $k' \in [\underline{k}, +\infty)$ respectively. In particular, neither depends on the permanent type j.

Similarly, the flow utility for j, $U^j(c,G) = \theta^j \log(e^j c) + (1-\theta^j)\log(G)$, can be equivalently written as

$$U^j(c,G) = \theta^j \log(c) + (1-\theta^j)\log(G), \tag{12}$$

where we drop the irrelevant additive constant $\theta^j \log(e^j)$. $U^j(c,G)$ in (12) is identical to $U(c,G)$ in (6) except for θ^j replacing θ, i.e., the permanent type j affects individual preferences in terms of normalized variables (c, k') and policy G only to the extent that it affects the relative weight

[15] For expositional reasons we use two types, but the argument goes through with any finite number of types.

of the utility from private consumption.[16] Since θ^j only scales life-time utility, the permanent type j does not affect the household's Euler equation and the decision rules for $\left(c, k'\right)$ for any given k and G. Of course, even with the same choice of $\left(c, k'\right)$, actual consumption and capital holdings for a type j household, $(\tilde{c}^j, \tilde{k}'^j)$, is proportional to her e^j and differs across $j = H$ and $j = L$ by a constant scale (e^H / e^L).

The economy, including the production technology and the political decision mechanism, is otherwise unchanged. In particular, a type-j household with wealth $\tilde{k}^j = e^j k$ has the weight $(e^j k^+)^\chi = (e^j)^\chi (k^+)^\chi$ in the social welfare function, i.e., a type with a higher e^j enjoys a higher political power on average when $\chi > 0$.

The following proposition states a "politico-economic aggregation across permanent types"-result: the new economy with permanent type differences shares the same time path (hence a fortiori the same comovement properties) of aggregate variables as our benchmark economy with a particular homogeneous preference parameter.

Proposition For any given wealth bias χ, consider an economy with *ex ante* heterogeneity defined by $\{e^j, \theta^j, \phi^j\}_{j \in \mathbb{J}}$, and an economy defined as in Section 3.1-3.3 with $\theta = (\sum_j (e^j)^\chi \phi^j \theta^j / \sum_j (e^j)^\chi \phi^j)$. If the initial joint distribution for the households in type j is generated as $(\tilde{k}^j = e^j k, \epsilon, \beta)$, where (k, ϵ, β) is drawn from the initial distribution $\Gamma_0(k, \epsilon, \beta)$ for the economy with no ex ante heterogeneity,[17] and if the realization of aggregate shocks $(z_t)_{t=0}^\infty$ is the same, then they have the same Markov-perfect equilibrium path for aggregate variables $(C_t, K_t, G_t)_{t=0}^\infty$.

The macroeconomic aggregates in the new economy with heterogeneous permanent types can thus be mapped to our benchmark economy with homogeneous types. In addition, the effect of *ex ante* heterogeneity matters only through the fact that it introduces a level effect on the preference parameter θ: the average ideology θ is closer to θ^H with a higher permanent difference e^H and group size ϕ^H, and a higher pro-wealth bias $\chi > 0$. Although such a level effect on θ is important for understanding the long-run steady state of the economy, it does not necessarily affect its fluctuations. Furthermore, to the extent that the economy with ex ante heterogeneity has to match the same average G/Y ratio for each given χ as the baseline economy with ex ante homogenous households, the calibration of $\{e^j, \theta^j, \phi^j\}_{j \in \mathbb{J}}$ has to lead to

[16]We note that this observation does not depend on our particular assumption of log utility. This property holds as long as life-time utility is separable in (c, G) and homothetic in c.

[17]Notice that given the aforementioned normalization this assumption makes the initial aggregate capital stock and the aggregate labor input the same in both economies. This assumption about the initial distribution is only needed to state the proposition in a strong form with identical aggregate paths. If the economy starts from an arbitrary initial distribution and the model features a global convergence property in terms of cross-sectional distributions, then the specification of the initial distribution is not important for the long-run stochastic behavior of the economy, in particular it is unimportant for its comovement properties. We verified with numerical simulations that indeed the initial distribution is irrelevant for our results.

the same θ as in the economy with no ex ante heterogeneity.

Going beyond this particular result, we believe that the general observation, i.e., permanent heterogeneity affects rather the long-run steady state of the economy as opposed to its fluctuations, is likely at play quantitatively even when an exact politico-economic aggregation result does not hold, as long as the permanent differences change at a much lower frequency than the fluctuations that are the focus of the analysis. Put differently, if agents increased their θ with their income or wealth over the medium-term frequency that we study in this paper, then our decoupling mechanism would indeed be counteracted.[18]

The proof of the proposition follows from two observations. First, the two economies share the same state variables and decision rules as a function of both state variables and (Ψ, H). Recall that in the economy with $\{e^j, \theta^j, \phi^j\}_{j \in \mathbb{J}}$, the decision rule for (c, k') does not depend on j and θ^j. As a result, the period-t distribution in (k, ϵ, β) within each j, $\Gamma_t(k, \epsilon, \beta)$, does not depend on j, hence is a sufficient state variable to summarize the cross-sectional distribution, provided that the initial distribution is equal to the same $\Gamma_0(k, \epsilon, \beta)$ for each j.

Second, the two economies share the same weighted social welfare, viewed as a function of $(\Gamma, z, G; \Psi, H)$. To see this, note that the permanent type $j = L$ and $j = H$ with the same k share the same decision rule for (c, k') and differ only by their political weight $(e^j)^\chi (k^+)^\chi$ and the relative weight (θ^j) in the flow utility $U^j(c, G)$. As a result, a weighted sum across j for all households with normalized capital stock k leads to a weighted flow utility for these households with capital k of $\left(k^+\right)^\chi \sum_j (e^j)^\chi \phi^j U^j(c, G)$, which is equivalent (up to an irrelevant multiplicative constant) to $\left(k^+\right)^\chi \left(\theta \log(c) + (1-\theta)\log(G)\right)$ with $\theta = (\sum_j (e^j)^\chi \phi^j \theta^j / \sum_j (e^j)^\chi \phi^j)$. Given that they share the same flow utility in each period, these two economies also share the same social welfare function defined in terms of life-time utility. Combining both observations, we can infer that the two economies share the same Markov-perfect equilibrium functions in terms of $\Psi(\Gamma, z)$ and decision rules, hence the same equilibrium path in terms of aggregate variables.

[18]In a similar vein, we verified numerically that assuming the same borrowing constraint across types does not have quantitatively large effects, i.e. the two economies are still very similar, although in this case the proposition does not hold exactly anymore.

C Computational Algorithm - Appendix

Algorithm 1 *Fixed Point Iteration on* (H,Ψ)

Step 0: Select a set of summary statistics of the wealth distribution $(K, Gini(k))$ *and fix the functional form. Start from an initial guess of coefficients* $\{a_0^0,\ldots,a_4^0\},\{\tilde{a}_0^0,\ldots,\tilde{a}_4^0\},\{b_0^0,\ldots,b_2^0\}$ *to get initial conjectured functions* $\left(H^0,\Psi^0\right)$*. Set up a convergence criterion* ε*.*

Step 1: In step n, imposing (H^n,Ψ^n) *in the best-response optimization problem, use value function iteration to solve the household's parametric dynamic programming problem. Get the continuation value function* $v^n\left(k,\epsilon,\beta,\Gamma,z;\Psi^n,H^n\right)$*.*

Step 2: Without imposing Ψ^n *and instead varying G freely on a finite grid, use* H^n *and* $v^n\left(k,\epsilon,\beta,\Gamma,z;\Psi^n,H^n\right)$ *to solve for the best-response value function* $J^n\left(k,\epsilon,\beta,\Gamma,z,G;\Psi^n,H^n\right)$ *and decision rule* $h^n\left(k,\epsilon,\beta,\Gamma,z,G;\Psi^n,H^n\right)$*.*

Step 3: Simulate the economy using N_H *households and T periods. In each period t of the simulation, calculate the equilibrium policy* $G_t^{eq.}$ *using* $J^n\left(k,\epsilon,\beta,\Gamma,z,G;\Psi^n,H^n\right)$ *and the social choice rule. Calculate the best response decision based on* $h^n\left(k,\epsilon,\beta,\Gamma,z,G;\Psi^n,H^n\right)$ *for both equilibrium* $G_t^{eq.}$ *and pre-specified* N_G *grid points of G,* $\left(G_{t,i}\right)_{i=1}^{N_G}$*. Gather a time series of* $\left(K_{t+1}^{eq.},\left(K_{t+1,i}\right)_{i=1}^{N_G},Gini_{t+1}^{eq.},\left(Gini_{t+1,i}\right)_{i=1}^{N_G},G_t^{eq.},\left(G_{t,i}\right)_{i=1}^{N_G}\right)_{t=1}^{T}$*, i.e. capital statistics both on* $(K_{t+1}^{eq.})$ *and off-equilibrium path* $\left(\left(K_{t+1,i}\right)_{i=1}^{N_G}\right)$*, with a total sample size of* $T(1+N_G)$*.*

Step 4: Use gathered time series to get – separately for each value of the z-grid – OLS estimates of $\{\widehat{a}_0^n,\ldots,\widehat{a}_4^n\},\{\widehat{\tilde{a}}_0^n,\ldots,\widehat{\tilde{a}}_4^n\},\{\widehat{b}_0^n,\ldots,\widehat{b}_2^n\}$*, which with a slight abuse of notation we summarize as* $\left(\widehat{H}^n,\widehat{\Psi}^n\right)$*. Notice that obviously* $\widehat{H}^n$ *is updated on both the on- and off-equilibrium paths,* $\widehat{\Psi}^n$ *only on the on-equilibrium path.*

Step 5: If $|H^n-\widehat{H}^n|<\varepsilon$ *and* $|\Psi^n-\widehat{\Psi}^n|<\varepsilon$*, stop. Otherwise, set*

$$
\begin{aligned}
H^{n+1} &= \alpha_H\times\widehat{H}^n+(1-\alpha_H)\times H^n,\\
\Psi^{n+1} &= \alpha_\Psi\times\widehat{\Psi}^n+(1-\alpha_\Psi)\times\Psi^n,
\end{aligned}
$$

with $\alpha_H,\alpha_\Psi\in(0,1]$*, and go to step 1.*

Step 6: Check whether the R2 of the final OLS regressions are high enough to convey confidence that the true equilibrium rule is well approximated. Otherwise go to step 0. [19].

[19] We chose $\varepsilon=10^{-4}$, $N_H=60{,}000$, $T=1{,}500$, of which we discard the first 500, when we update the KS-rules or compute summary statistics. Following Krusell and Smith (1998), we also make sure that these 60,000 agents are always distributed according to the stationary distributions of the Markov chains that govern ϵ and β, and thus avoid introducing artificial aggregate uncertainty owing to the small deviation from the law of large numbers. To eliminate sampling error, we use the same series of aggregate shocks for all iterations and all model simulations.

D Calibration - Appendix

Table 5: COMMON PARAMETERS

Parameter	δ	α	$\underline{k}$
Value	0.1	0.36	-0.01

Table 6: MARKOV CHAIN: AGGREGATE PRODUCTIVITY

State	z_1	z_2	z_3	z_4	z_5
Value	0.9182	0.9582	1	1.0436	1.0891
z_1	0.6306	0.3676	0.0018	0.0000	0.0000
z_2	0.0382	0.7538	0.2077	0.0003	0.0000
z_3	0.0001	0.0980	0.8039	0.0980	0.0001
z_4	0.0000	0.0003	0.2077	0.7538	0.0382
z_5	0.0000	0.0000	0.0018	0.3676	0.6306

Notes: This Markov chain is based on an autocorrelation coefficient of 0.8145 and conditional standard deviation of 0.0165. It was generated with Tauchen's (see Tauchen, 1986) discretization method and a width-parameter of 3.

Table 7: MARKOV CHAIN: IDIOSYNCRATIC LABOR PRODUCTIVITY

State	ϵ_1	ϵ_2	ϵ_3	ϵ_4	ϵ_5	ϵ_6	ϵ_7	ϵ_8	ϵ_9
Value	0.4420	0.5421	0.6648	0.8154	1	1.2264	1.5041	1.8447	2.2623
ϵ_1	0.2854	0.4292	0.2409	0.0422	0.0023	0.0000	0.0000	0.0000	0.0000
ϵ_2	0.0782	0.3102	0.4140	0.1739	0.0227	0.0009	0.0000	0.0000	0.0000
ϵ_3	0.0117	0.1167	0.3716	0.3716	0.1167	0.0113	0.0003	0.0000	0.0000
ϵ_4	0.0009	0.0227	0.1739	0.4140	0.3102	0.0728	0.0053	0.0001	0.0000
ϵ_5	0.0000	0.0023	0.0422	0.2409	0.4292	0.2409	0.0422	0.0023	0.0000
ϵ_6	0.0000	0.0001	0.0053	0.0728	0.3102	0.4140	0.1739	0.0227	0.0009
ϵ_7	0.0000	0.0000	0.0003	0.0113	0.1167	0.3716	0.3716	0.1167	0.0117
ϵ_8	0.0000	0.0000	0.0000	0.0009	0.0227	0.1739	0.4140	0.3102	0.0782
ϵ_9	0.0000	0.0000	0.0000	0.0000	0.0023	0.0422	0.2409	0.4292	0.2854

Notes: This Markov chain is based on an autocorrelation coefficient of 0.75 and conditional standard deviation of 0.18. It was generated with Tauchen's (see Tauchen, 1986) discretization method and a width-parameter of 3.

Table 8: MARKOV CHAIN: DISCOUNT FACTOR

State	β_1	β_2	β_3
Value	0.94	0.96	0.98
β_1	0.9800	0.0200	0
β_2	0.0025	0.9950	0.0025
β_3	0	0.0200	0.9800

The calibration of θ***:***

Baseline case and its three variants: $[\chi = 0, \theta = 0.78]$, $[\chi = 0.1, \theta = 0.79]$, $[\chi = 0.2, \theta = 0.795]$, $[\chi = 0.3, \theta = 0.805]$, $[\chi = 0.4, \theta = 0.81]$, $[\chi = 0.5, \theta = 0.815]$, $[\chi = 0.55, \theta = 0.82]$, $[\chi = 0.6, \theta = 0.82]$,$[\chi = 0.7, \theta = 0.83]$, $[\chi = 0.8, \theta = 0.84]$,$[\chi = 0.9, \theta = 0.85]$,$[\chi = 1.0, \theta = 0.86]$.

Representative agent case: $\theta = 0.78$.

'Majority voting'-case and its three variants: $[\chi = 0, \theta = 0.76]$, $[\chi = 0.1, \theta = 0.765]$, $[\chi = 0.2, \theta = 0.77]$, $[\chi = 0.3, \theta = 0.775]$, $[\chi = 0.4, \theta = 0.78]$, $[\chi = 0.5, \theta = 0.785]$, $[\chi = 0.6, \theta = 0.795]$,$[\chi = 0.7, \theta = 0.805]$,$[\chi = 0.75, \theta = 0.815]$,$[\chi = 0.79, \theta = 0.82]$,$[\chi = 0.8, \theta = 0.82]$,$[\chi = 0.9, \theta = 0.835]$,$[\chi = 1.0, \theta = 0.85]$.

The calibration of the Domeij-Heathcote cases:

Table 9: PARAMETERS FOR THE MODELS IN TABLE 4

Model	Column 3			Column 4			Column 5		
Parameter	χ	θ	$\underline{k}$	χ	θ	$\underline{k}$	χ	θ	$\underline{k}$
Value	0.40	0.815	-0.405	0.40	0.81	-0.0009	0.40	0.805	0

Table 10: MARKOV CHAIN: IDIOSYNCRATIC LABOR PRODUCTIVITY PROCESSES FOR THE MODELS IN TABLE 4

Model	Column 3			Column 4			Column 5		
State	ϵ_1	ϵ_2	ϵ_3	ϵ_1	ϵ_2	ϵ_3	ϵ_1	ϵ_2	ϵ_3
Value	0.0473	1.0080	5.6911	0.5292	0.9750	4.1769	0.1978	0.9975	6.1467
ϵ_1	0.7502	0.2498	0	0.7521	0.2479	0	0.7500	0.2500	0
ϵ_2	0.0015	0.9970	0.0015	0.0079	0.9842	0.0079	0.0032	0.9936	0.0032
ϵ_3	0	0.2498	0.7502	0	0.2479	0.7521	0	0.2500	0.7500

Notes: This Markov chain is based on an autocorrelation coefficient of 0.75 and conditional standard deviation of 0.18. It was generated with the Domeij-Heathcote discretization method (see Domeij and Heathcote, 2004).

E Data - Appendix

Table 11: COMOVEMENT AND PERSISTENCE FOR DISAGGREGATE GOVERNMENT PURCHASES - DISAGGREGATION ACCORDING TO ADMINISTRATIVE UNITS

Moment	Correl. w. Y	Correl. w. Y-Lag.	Correl. w. C	Autocorrel. 1st-order	Frac. of G
G	0.35	0.51	0.35	0.79	100.0%
GC	0.26	0.43	0.30	0.78	84.6%
GI	0.47	0.59	0.39	0.75	15.6%
GND	0.47	0.58	0.49	0.74	68.8%
$GNDC$	0.19	0.34	0.38	0.70	56.5%
$GNDI$	0.60	0.66	0.47	0.74	12.0%
GF	0.15	0.27	0.11	0.82	42.7%
GFC	0.12	0.25	0.09	0.82	38.3%
GFI	0.18	0.28	0.10	0.73	5.0%
GFD	0.13	0.26	0.13	0.83	31.2%
$GFDC$	0.13	0.24	0.12	0.85	28.0%
$GFDI$	0.08	0.27	0.05	0.71	3.6%
$GFND$	0.04	-0.01	-0.12	0.63	11.7%
$GFNDC$	-0.09	-0.04	-0.20	0.47	10.3%
$GFNDI$	0.27	0.03	0.06	0.71	1.5%
GSL	0.49	0.65	0.59	0.76	57.1%
$GSLC$	0.24	0.39	0.48	0.77	46.3%
$GSLI$	0.62	0.73	0.52	0.72	10.7%
GS	0.40	0.65	0.47	0.72	18.4%
GSC	0.10	0.41	0.27	0.73	14.0%
GSI	0.57	0.62	0.55	0.69	4.2%
GL	0.49	0.58	0.58	0.70	38.7%
GLC	0.27	0.30	0.51	0.70	32.2%
GLI	0.52	0.65	0.39	0.70	6.5%

Notes: G denotes government consumption and gross investment expenditures, a C in an acronym means consumption, an I investment. D stands for defense spending, ND for nondefense. F means federal government, SL the aggregate of state and local governments. S stands for the state level and L for the local level. All variables are annual, they range from 1960-2006. Categories G until $GSLI$ are deflated by their corresponding deflators. For the separate state and local level data NIPA does not publish separate price indices. We therefore use the aggregate state and local price deflator for GS and GL, and the consumption- and investment-specific aggregate state and local price deflator for GSC, GLC and GSI, GLI, respectively. For columns 2-5 all data are logged and filtered with a Hodrick-Prescott filter with smoothing parameter 100. The last column shows the fraction of each component of government purchases in total G. Sources: Tables 3.9.4 and 3.9.5 from the NIPA accounts for G until $GSLI$. The separate data for the state and local level come from Tables 3.20 and 3.21 from the NIPA accounts.

Table 12: COMOVEMENT AND PERSISTENCE FOR DISAGGREGATE GOVERNMENT PURCHASES - FUNCTIONAL DISAGGREGATION

Moment	Correl. w. Y	Correl. w. Y-Lag.	Correl. w. C	Autocorrel. 1st-order	Frac. of G
General public service	0.10	0.17	0.38	0.69	8.32%
National defense	0.12	0.26	0.13	0.83	29.75%
Public order and safety	0.05	0.33	0.21	0.56	9.06%
Economic affairs	0.37	0.35	0.22	0.72	16.36%
Transportation	0.51	0.49	0.50	0.66	10.04%
Space	0.15	0.07	-0.06	0.83	1.10%
Other economic affairs	0.00	0.10	-0.08	0.28	5.17%
Housing & comm. serv.	0.22	0.47	0.13	0.41	2.30%
Health	-0.29	-0.03	-0.19	0.76	4.49%
Recreation and culture	0.03	0.37	0.19	0.61	1.36%
Education	0.51	0.58	0.58	0.79	25.47%
Income security	0.04	0.13	0.04	0.45	2.87%

Notes: Data source is the BEA (NIPA table 3.15.5). All variables are annual, the sample goes from 1960-2006. They are deflated by their corresponding deflators (NIPA table 3.15.4), and for columns 2-5 logged and filtered with a Hodrick-Prescott filter with smoothing parameter 100. 'Housing & comm. serv.' stands for 'Housing and community services'.

F Countercyclical Inequality of Total Available Resources - Appendix

Define the total available resource for each household who belongs to a group with a specific ϵ as

$$\begin{aligned}\omega(k,\epsilon;K,L,z) &= (1-\delta)k + r(K,L,z)k + w(K,L,z)\widetilde{l}\epsilon \\ &= (1-\delta)k + z\alpha\left(\frac{K}{L}\right)^{\alpha-1}k + z(1-\alpha)\left(\frac{K}{L}\right)^{\alpha}\widetilde{l}\epsilon \\ &= z\left[\left(\frac{1-\delta}{z} + \alpha\left(\frac{K}{L}\right)^{\alpha-1}\right)k + (1-\alpha)\left(\frac{K}{L}\right)^{\alpha}\widetilde{l}\epsilon\right].\end{aligned}$$

Notice that a higher z introduces two effects on ω: a level effect through increasing income, as well as a relative inequality effect by changing the relative weight between the gross return of capital and labor.

Define the Lorenz curve in ω as (for the sake of readability we leave out the function arguments ϵ, K and L):

$$L(x;z) = \frac{\int_0^x \omega(k(i);z)\,di}{\int_0^1 \omega(k(i);z)\,di}.$$

The following result states that inequality in total available resources, ω, is decreasing in z.

Proposition Let $\omega(k,\epsilon;K,L,z) > 0$, then for every $0 < x < 1$, $L(x;z_2) > L(x;z_1)$ whenever $z_2 > z_1$.

Proof. The result follows from standard arguments. To start with, we know that

$$\begin{aligned}& L(x;z_1) - L(x;z_2) \\ =\; & \int_0^x \left(\frac{\omega(k(j);z_1)}{\int_0^1 \omega(k(i);z_1)\,di} - \frac{\omega(k(j);z_2)}{\int_0^1 \omega(k(i);z_2)\,di}\right)dj \\ =\; & \int_0^x \left[\frac{\omega(k(j);z_2)}{\int_0^1 \omega(k(i);z_1)\,di}\left(\frac{\omega(k(j);z_1)}{\omega(k(j);z_2)} - \frac{\int_0^1 \omega(k(i);z_1)\,di}{\int_0^1 \omega(k(i);z_2)\,di}\right)\right]dj.\end{aligned}$$

To proceed, notice that

$$\begin{aligned}
\omega(k; z_2) &= (1-\delta)k + z_2\alpha\left(\frac{K}{L}\right)^{\alpha-1} k + z_2(1-\alpha)\left(\frac{K}{L}\right)^{\alpha}\widetilde{l}\epsilon \\
&= \left[\begin{array}{c} (1-\delta)k + z_1\alpha\left(\frac{K}{L}\right)^{\alpha-1} k + z_1(1-\alpha)\left(\frac{K}{L}\right)^{\alpha}\widetilde{l}\epsilon \\ +(z_2-z_1)\left[\alpha\left(\frac{K}{L}\right)^{\alpha-1} k + (1-\alpha)\left(\frac{K}{L}\right)^{\alpha}\widetilde{l}\epsilon\right] \end{array}\right] \\
&= \omega(k; z_1) + (z_2-z_1)\left[\alpha\left(\frac{K}{L}\right)^{\alpha-1} k + (1-\alpha)\left(\frac{K}{L}\right)^{\alpha}\widetilde{l}\epsilon\right]
\end{aligned}$$

so that

$$\frac{\omega(k; z_1)}{\omega(k; z_2)} = 1 - \frac{(z_2-z_1)\left[\alpha\left(\frac{K}{L}\right)^{\alpha-1} k + (1-\alpha)\left(\frac{K}{L}\right)^{\alpha}\widetilde{l}\epsilon\right]}{(1-\delta)k + z_2\alpha\left(\frac{K}{L}\right)^{\alpha-1} k + z_2(1-\alpha)\left(\frac{K}{L}\right)^{\alpha}\widetilde{l}\epsilon},$$

which is a strictly increasing function in k. From this monotonicity, $\frac{\omega(k(j); z_1)}{\omega(k(j); z_2)} - \frac{\int_0^1 \omega(k(i); z_1)di}{\int_0^1 \omega(k(i); z_2)di}$ is increasing in j, which implies that for $j > 0$ it crosses zero at most once and from below if crossing happens. Because $\frac{\omega(k(j); z_2)}{\int_0^1 \omega(k(i); z_1)di} > 0$ for $j > 0$, the same crossing property holds for the whole integrand. Next, we show that the integrand crosses zero exactly once and from below.

Suppose that the integrand does not cross zero and hence is either positive or negative for every $0 < j < 1$. Given $L(0; z_1) - L(0; z_2) = 0$, this immediately implies that $L(1; z_1) - L(1; z_2) > 0$ (< 0) if the integrand is positive (negative), which contradicts the fact that $L(1; z_1) - L(1; z_2) = 1 - 1 = 0$ from the definition of a Lorenz curve. As a result, the integrand actually crosses zero once and from below.

Combining this with the fact that $L(1; z_1) - L(1; z_2) = 0$, it must be true that $L(x; z_1) - L(x; z_2) < 0$ for every $0 < x < 1$. This finishes the proof. ■