Essays in Dynamic Macroeconomics and Political Economy

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Marco M. Sorge
aus Napoli

Bonn 2012

Dekan:	Prof. Dr. Klaus Sandmann
Erstreferent:	Prof. Dr. Martin F. Hellwig
Zweitreferent:	Prof. Dr. Thomas Hintermaier

Tag der mündlichen Prüfung: 20. Dezember 2011

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

To Lucy and Viola

Acknowledgements

This thesis presents the main results of my research at the Bonn Graduate School of Economics. I am indebted to a number of people who supported me and my work during the past years.

Firstly, I owe my most sincere gratitude to my Ph.D. advisors Prof. Martin F. Hellwig and Prof. Thomas Hintermaier for their excellent guidance and help whenever I needed, and for the freedom they granted me to develop my own research ideas. The numerous comments and the fruitful discussions we had on the topics dealt with in the present work, have been a constant source of inspiration. All errors, of course, are my own.

I am very thankful to my co-authors Giuseppe Albanese and Francesco Carravetta for the excellent cooperation, and to Prof. Norbert Christopeit and Dr. Michael P. Evers for their advice to tackle several computational issues.

My gratefulness goes to all the participants of the BGSE workshops (University of Bonn) and of the other conferences and meetings where some of the results reported here were first presented, for various remarks and suggestions, which helped to improve my work.

Moreover, I am grateful to Prof. Urs Schweizer for his incessant effort in managing the doctoral program in Quantitative Economics at the BGSE.

Most of all I want to thank my family for their endless patience, and my little lovely star Lucianna, for her unbreakable faith that this day would have come. Words fail to express my gratitude.

Contents

List of Figures

Introduction

This dissertation attempts in four essays to contribute to economic research in two separate fields, and is therefore divided into two self-contained parts.

Part I (Dynamic Macroeconomics) is rooted in the modern macroeconomic literature that aims at exploring theoretical departures from the classic rational expectations (RE) paradigm. According to the latter, economic agents know at all dates the full state of the economy as well as the structure which generated the state itself. Our study rather accounts for a time-varying structure and imperfect observability of the model's variables (chapter 1), and for advanced information on the future states - that is, for news shocks and anticipation (chapter 2).

In chapter 1[1], the problem of finding a solution to time-varying linear RE systems involving past expectations of the future state values and noisy observations is addressed. This is done by introducing a novel dynamic policy optimization formulation for forward-looking models, consistent with the classical rational choice paradigm. For a well-identified class of time-dependent RE frameworks, it is shown that there exists always an equilibrium path having the property of being the closest, in mean square, to the state motion of the autoregressive dynamic equation governing the perfect foresight behavior of the economic system. A recursive algorithm - based upon Kalman filtering - providing the exact expression for the conditional expectations (hence, the solution) and the optimal filtering estimate, is also presented.

The approach developed in chapter 1 is able to handle higher order models as well as nonlinear ones. An extension to the relevant class of Markov-switching RE models, which have recently been advocated to investigate the role of regime switching monetary policy in New-Keynesian frameworks and the effects of parameter instability (e.g., Davig and Leeper, 2007; Liu *et al.*, 2009), is also presented.

Most importantly, the solution algorithm of chapter 1 naturally applies to settings with time-dependent structures, which is attractive for empirical applications. Classical methods for solving linear stochastic systems under RE, such as Blanchard and Kahn (1980), King and Watson (1998, 2002), Klein (2000) and Sims (2002), postulate indeed a structural representation in which the parameters that govern the

[1]This chapter is based on Carravetta and Sorge (2010, 2011).

behavior of the dynamic system do not vary over time[2]. From a different perspective, the analysis of the combination of stochastic control techniques, time-varying parameters and the RE hypothesis is part of a promising research project aiming at mitigating the well-known Lucas critique (1976) by providing methods able to yield optimal policy instruments in models under RE, even in the presence of time-dependent - possibly uncertain - model structure (e.g. Amman and Kendrick, 2003; Tucci, 2004). We build upon the mentioned strands of literature by proposing a Kalman filter-based technique for the estimation of the expectational term forcing a general vector stochastic system with time-varying structure[3], for which initial conditions knowledge only is required.

In chapter 2[4], we study identification of linear dynamic RE models under news shocks. The main question addressed is whether these models are empirically distinguishable from lagged expectations RE systems under i.i.d. fundamental shocks which rather allow for equilibrium serially correlated sunspots.

Moving from the seminal work of Azariadis (1981) and Kass and Shell (1983), economic theorists have paid attention to the role of extrinsic uncertainty as driver of changes in economic beliefs, which may involve causal effects on cyclical fluctuations aside from shocks to fundamentals (e.g. Schmitt-Grohé, 1997; Benhabib and Farmer, 1999). A related though different strand of literature has recently grown up which exploits the idea that advance information or "news" about future developments in the economy can induce cycles in the major economic aggregates, regardless of whether the information content of the news is later rectified or not (e.g. Beaudry and Portier, 2004; Jaimovich and Rebelo, 2009).

Chapter 2 explores the empirical separability of news shocks and sunspots models. By means of the general martingale solution approach, we show that it may prove impossible to decide on an econometric basis whether the actually observed data is generated by determinate models driven by news shocks or rather by indeterminate ones forced by sunspot variables. More specifically, for any exactly identified news shocks model, there exists an observationally equivalent class of indeterminate RE systems, which are only subject to i.i.d. fundamental shocks. Since the alterna-

[2]An important exception, though limited to a specific case, is Blake (1990), which extends the Blanchard and Kahn (1980)'s famous framework to one where coefficients are policy-dependent.

[3]For example, the inflation target or the parameters governing the reaction of the interest rate to current values of inflation and output within a Taylor-type rule for the baseline New Keynesian framework (e.g. Galì, 2008).

[4]This chapter is based on Sorge (2011a).

tive models possess different determinacy properties, different implications for policy making are also likely to arise. Given the prominent position gained by theories regarding news shocks and anticipation effects in the modern business cycle debate, this finding points to the opportunity of supplementing likelihood-based empirical investigations of news shocks in estimated DSGE models with testing strategies for the indeterminacy hypothesis.

Part II of the dissertation (Political Economy) aims to contribute to the literature on the role of self-interested groups in the political arena.

In chapter 3[5], we investigate theoretically how the presence of (corruptible) judiciaries that oversee the political process impacts on one of the mechanisms by which lobby groups can influence policy outcomes, i.e. bribery. Since the scope of corruption depends on the expected benefits to politicians from becoming corrupt, systematic and well-targeted efforts toward the investigation and prosecution of bribery cases might serve as a powerful device for corruption deterrence. However, effective judicial oversight typically faces institutional and operational constraints, like political interference or judicial subversion, which may hinder the proper functioning of this mechanism. In fact, independent judiciaries that act in corrupt societies are vulnerable to taking bribes (Glaeser and Shleifer, 2002). We develop an endogenous policy model to shed some light on these and other related issues concerning political corruption and judicial independence.

We show that judicial independence is a necessary condition for deterrence effects to arise from the oversight activity of judiciaries. In fact, dependent judges are not able to prevent the interest group and the government from maximizing the profits from the deals between them. From an institutional design perspective, our results make a strong case for insulating judicial branches from political interference.

While the independence of the judiciary is crucial to its effectiveness, it may also foster corruption in this branch because no other government entity has the authority to oversee it. Hence, judges must be subject to mechanisms that hold them accountable for their institutional role. Our analysis suggests that preserving the efficiency of independent judiciaries can serve as an instrument for self-enforced judicial accountability, even in the presence of corrupt judges.

[5]This chapter is based on Albanese and Sorge (2011).

In chapter 4[6], we study the process of legislative delegation in the presence of bureaucratic lobbying. Strategic appointments are introduced in Bennedsen and Feldmann (2006)'s delegation model in which a legislator delegates policy authority to a bureaucracy, who holds superior information regarding the political environment. An organized group is able to influence the process by initiating bargaining in policy implementation. The bureaucracy is not viewed as an exogenous entity, but rather assumed to be endogenously appointed by the government administration.

We show that the possibility of strategic agency selection fully restores general results from the conventional theory of delegation. In particular, bureaucratic lobbying never reduces the scope of delegation across different political systems (parliamentary versus separation of powers), as it engenders no influence on the extent of (expected) policy bias induced by delegated legislation. By contrast, the optimal degree of delegated authority emerges as an exclusive relationship between agency discretion, on the one hand, and the ideological conflict between the higher-level institutions (the legislature and the administration), and the uncertainty surrounding the political environment, on the other.

The model has important consequences for the theory of agenda setting and political control. First, our analysis contributes to the well-known debate over devices available to a legislature to control bureaucratic policy-making, for it suggests that strategic appointments may work as a substitute for legislative oversight (e.g., Gailmard, 2009). An important corollary of this result is that the legislature need not shape delegated legislation on the degree of interest group influence on agency decision-making. Second, from the perspective of optimal statutory design, the model predicts that divided governments should be characterized by more stringent boundaries to which agency decisions ought to conform, independently of the active participation of interest groups in agency decision-making.

Although collected into two parts, the next four chapters each present one idea as a self-contained unit.

[6]This chapter is based on Sorge (2011b).

Part I

Dynamic Macroeconomics

Chapter 1

Kalman Filter Approach to Solution of Noisily Observed Rational Expectations Models

1.1 Introduction

Since the early work of Muth (1961) and Lucas (1972, 1976), the concept of rational expectations (RE) has become the standard tool of modeling expectations in dynamic stochastic macroeconomic models. It essentially reduces to the assumption that agents collect and make optimal use of all available (pertinent) information as to the economic environment when formulating their forecasts of economic variables of interest (e.g., prices, interest rates, government policies). More specifically, the RE hypothesis requires that the prediction made by the forecaster, conditional on all the information available at the time the prediction is made, be consistent with the forecast model derived from the underlying economic structure, where no systematic forecast errors are involved (e.g., Shiller, 1978; Barth *et al.*, 1982).

In models where expectations on future states influence current dynamics, an infinity of solutions under the assumption of RE has been proven to exist. Since this result was established by the early work of Sargent and Wallace (1973), Shiller (1978) and Blanchard (1979), several attempts have been made toward an evaluation of the relevance of this issue to macroeconomic modeling - e.g., McCallum (1983) - while other authors have aimed at offering a wide array of instruments intended to select a particular solution among all the available ones. Such non-uniqueness property, arising even in linear dynamic systems, has thrown into question the rationality of equilibria and somewhat been interpreted as a serious weakness for the RE hypothesis.

A more subtle issue, on the ground of which the latter has been often criticized, lies in that, even in its strongest forms, it fails to shed light on the process by which economic agents translate current information - what they truly observe and know about - into optimal forecasts, which are simply assumed - and turn to be self-fulfilled - not systematically different from equilibrium outcomes (Lucas and Prescott, 1971).

Attempting to address such issues, Basar (1989) and De Santis *et al.* (1993) consider a slightly different theoretical setting. In the former, an alternative formulation of RE models is proposed as scalar linear difference stochastic equations forced by a function of the information available up to the current time, which is determined by solving a quadratic cost index optimization problem over a finite time interval. Existence, uniqueness and minimum variance property of the state solution for the policy optimization formulation are derived and compared with the direct solution to the RE dynamic model[1].

An analogous problem statement, generalized to the vectorial case, is considered in De Santis *et al.* (1993), where the concept of "ideal" behavior of a RE model is introduced and a method for estimating the control function based on a reference framework adaptive algorithm is proposed. The reference model is chosen as the autoregressive stochastic difference equation describing the ideal situation of perfect foresight and the optimal[2] estimate of the future (two-step ahead) state is computed via the Kalman filter (Kalman, 1960). An approximate solution of the actual problem is then determined by using the output (noisy) measurements of the actual system in the previously obtained estimator and then using the latter as control function.

Following De Santis *et al.* (1993) and Carravetta (1996), in this work we address the problem of finding the solution of the general RE model in state-space form[3], involving past expectations of the future state values and noisy observations of the state vector. For a well-specified subset in the class of RE models dealt with in present work, a "nearly ideal" RE solution under a well-specified model reference adaptive technique and an estimation algorithm for it, based upon Kalman filtering, are provided. We show that the state motion is the closest, in mean square, to the evolution the system would have if agents were able to form an *exact* prediction of - i.e., perfectly foresee - the future state values.

[1]Similarly, in Chow (1980b) a general difference model forced by an economic agent's action is considered. Here the agent is assumed to maximize an objective expressed as a quadratic cost function and the problem of optimal estimation for linear RE models is dealt with, providing a family of consistent estimators. From a slightly different perspective, Chow (1980a) deals with structural RE models, where optimal control rules are applied to the reduced form equations with policy invariant coefficients.

[2]The criterion of optimality here refers to the minimum variance principle.

[3]The utility of state-space form representations for macroeconometric models is highlighted in Wall (1980). In this setting, estimation and model identification problems can be carried out for all the main expectations formation mechanisms by exploiting the recursive estimation methods for state-space models derived from Kalman filtering theory.

The basic structure of the procedure employed herein for combining econometric tools with dynamic economic theory builds upon Carravetta (1996). In this respect, the contribution of the present work is twofold. On the one hand, we show that what was implicitly conjectured in De Santis *et al.* (1993) and further argued in Carravetta (1996) as to the existence of an *exact* RE equilibrium, which could be found by means of a *causal* model forced by the optimal prediction of the perfect foresight state, is not factually correct, as it fails to obtain as a general property of RE models. A full characterization of RE models in the class we deal with is provided for which an RE equilibrium, equivalent to the nearly ideal state motion, exists as the unique (to within stochastic equivalence) solution of linear time-varying RE models amenable with Kalman filtering theory. On the other, we improve upon the previous results in the literature by constructing an estimation algorithm for linear models which exploits the actual measurement of the state vector, rather than being based on the fictitious output generated by the ideal (reference) model, as in De Santis *et al.* (1993).

The analysis presented here is related to different lines of research. Previous studies have focused on the fundamental problem in estimating structural RE models which stems from the presence of unobservable components (e.g., Chow, 1980b; Hansen and Sargent, 1980; Pesaran, 1989). In particular, Broze and Szafarz (1991) describe classical estimation schemes based on replacing the expectation term by suitable proxies (OLS estimator) or by the realized (observed) values of the variables (error-in-variable approach). In assessing the non-uniqueness specificity of (most) RE models, they offer a thorough analysis of the conditions under which the equilibrium path actually followed by the economy can be statistically determined *a posteriori* by using the data to estimate the unrestricted parameters of the general reduced form ("letting the data *ex post* select the right model"). Similarly, Tucci (2004) emphasizes the possibility of exploiting the observational equivalence between the linear (scalar) stationary RE models estimated via the error-in-variable method and an error-in-variable model with restricted time-varying parameters, in order to (indirectly) test for the RE hypothesis using estimated residuals. The approach followed by Broze and Szafarz (1991) and extended in Tucci (2004) differs from ours in that we rather question the modeling of the mechanism of RE formation from an ex ante (solution) point of view, according to which rational economic agents are likely to simulate econometricians when forming their expectations about the future state of the economy they act in ("letting the data - as they become available - help them select the optimal model").

From this perspective, our analysis is related to studies on the process of expectation formation. Previous work on this subject differs from the present one in that it generally focuses on learning behavior, that is the way systematic forecasting biases are eliminated over time (e.g., Marcet and Sargent, 1989; Evans and Honkapohja, 2001). Specifically, the adaptive learning literature endows boundedly rational agents with a forecasting model - the perceived law of motion of the economy - which can be an arbitrary function of past endogenous and past and current exogenous variables, and has to be optimally parameterized based on new data and observable (past) forecast errors. RE equilibria are thus regarded as asymptotic outcomes of this learning process, whenever conditions for convergence of agents' beliefs to the equilibrium values hold. Though methodologically related, our method also differs from the Bayesian learning literature (e.g., McGough, 2003; Bullard and Suda, 2008), as these studies typically assume that agents employ filtering techniques to update estimates of (possibly time-varying) parameters within not fully rational forecasting functions. Rather, our approach posits that non-fully rational agents may be thought of as revising their (best) estimate of the (hidden) variables governing the dynamics of the economic system as new observations are generated, when only a reduced information set - the measurement process - is available to them.

It is well known that the dimension of the solution set for RE models is closely related the stability properties of the latter, and that stability restrictions can be advocated in order to weaken the multiplicity issue (e.g., Blanchard and Kahn, 1980; Salemi, 1986). However, as the agents' expectations in RE frameworks are typically obtained by recursively iterating the system into the future, (asymptotic) stationarity is needed for this process to be well-defined. While equilibrium stability is usually enforced by the existence of transversality conditions in the underlying (infinite horizon) dynamic economic frameworks, there exist models for which no such boundary conditions arise or rather, though present, they do not serve as necessary optimality requirements (e.g., Halkin, 1974; Driskill, 2006). In this respect, we emphasize that, by providing a readily computable expression of the RE component both in finite and infinite horizon model representations, our method need not invoke approximation hypotheses or stability concepts to solve forward the system, for only initial conditions knowledge is required.

Our work also relates to the application of classical stochastic control theory techniques to RE dynamic systems. Starting with the seminal contributions of Lucas (1972) and Kydland and Prescott (1977), much criticism has been raised as to the use of control theory in economics, one of the major drawbacks being its inability to deal with RE due to the failure of the causality hypothesis (Aoki and Canzoneri, 1979; Chow, 1980a; Kendrick, 1981). The literature in the field has mostly aimed at overcoming the time inconsistency issue by providing methods for deriving optimal (policy) instruments in models under RE, even in the presence of time-dependent (possibly uncertain) parameters (e.g., Amman and Kendrick, 2003); interestingly, Tucci (2004) also discusses the suitability of adaptive control techniques in the presence of time-varying parameters to control RE systems estimated by the error-in-variable method. In our framework, we interpret the choice of the input sequence in a dynamic policy optimization setting as being based on the expectations agents have about future values of the state vector, conditional on past and present (noisy) observations. As a consequence, the problem of deriving reduced form representations free of expectational terms in order to compute the admissible set of instruments (Amman and Kendrick, 1999), is irrelevant for our results.

Canonical methods for solving linear stochastic models under RE, such as Blanchard and Kahn (1980), King and Watson (1998) and Sims (2002), postulate a structural representation in which the parameters that govern the behavior of the dynamic system are taken to be time-invariant[4]. In contrast to traditional approaches, the Kalman filtering-based solution method we propose is also able to handle frameworks in which the model parameters are allowed to change over time, as is the case in RE equilibrium models whose structure is supposed to react to policy variables. Accordingly, it offers a potentially helpful tool for obtaining suitable descriptions of several important scenarios occurring in macroeconometrics, such as inflationary situations which are surrounded by a rapidly evolving environment. The stationary case is also presented and sufficient conditions are provided which assure the existence of a steady state solution and of a filtering state estimate with constant steady state error covariance matrix. Finally, we briefly discuss an extension of our

[4]An important exception is Blake (1990), which provides a set of conditions for recovering the reduced RE form for linear models under anticipated policy reversal.

method to Markov-switching RE models in the latter part of this work.

The remaining chapter is organized as follows. In section 1.2 the problem we deal with is formally stated for linear time-varying RE models. In section 1.3 we present particular solutions for a general dynamic stochastic difference systems, including the one generated when forcing the latter with the optimal (minimum variance) estimate of the perfect foresight (ideal) state, fed by the measurement of the actual state variables - the result derived in De Santis *et al.* (1993) is thus revised accordingly. In section 1.4, a characterization of the subset in the class of RE models dealt with in the present study, which admits these particular solutions, is established, whereas section 1.5 is devoted to the analysis of the stationary case. In section 1.6, we study the Markov-switching case. Section 1.7 concludes. For the sake of exposition, all the proofs and technical details are relegated to the Appendix.

1.2 The model and the problem statement

In this work we are interested in reconsidering linear RE models and their solutions from a point of view which, along the lines of Basar (1989), deviates to some extent from the approaches typically adopted in the macroeconomic literature. To illustrate this point, let us first consider the following linear stochastic discrete-time system:

$$x_{t+1} = A_t x_t + B_t u_t + v_t, \quad x_0 = \bar{x} \tag{1.1}$$

which generalizes the form presented in De Santis *et al.* (1993) to the time-varying parameters case. It stands also as a time-dependent version of the generalized expectations models (GEM) introduced in Wall (1980), to which many underlying econometric models can be transposed - e.g., all those described by autoregressive-moving average (ARMA) processes (Hamilton, 1994).

In equation (1.1), x_t, u_t, v_t are all vectors in $\Re^n$; A_t, B_t are real square $n \times n$ matrices, B_t non-singular for any t. The initial state $\bar{x}$ is assumed to be a zero-mean Gaussian random variable with known covariance matrix P_0. The sequences $\{x_t, u_t\}$ describe the dynamic evolution of an economic variable of interest and the action (control) a representative economic agent can exploit to influence the evolution of the economic system, respectively. The sequence $\{v_t\}$ represents an unobserved structural disturbance and is modeled as a zero-mean white Gaussian sequence with assigned positive definite covariance matrices $\{Q_t\}$.

We assume uncertainty in the measurement of the state variables (noisy observations) according to a linear transformation acting on the state process:

$$y_t = C_t x_t + w_t, \quad y_0 = \bar{y} \tag{1.2}$$

where $y_t \in \Re^m$ is the output vector, $C_t \in \Re^{m \times n}$ and the initial value $y_0 \equiv C_0 \bar{x} + w_0$ is denoted by $\bar{y}$; w_t is the zero-mean white Gaussian sequence with known covariance matrices $W_t \equiv E\left(w_t w_t^T\right)$, which accounts for the output measurement noise. For simplicity, the system-equation noises and the measurement errors $\{v_t\}$, $\{w_t\}$ are assumed to be mutually independent as well as independent of the initial state $\bar{x}$.

In the classical RE approach, the economic agent takes his decisions on the basis of his expectations about future values of the sequence $\{x_t\}$ - let us consider, for instance, the two-step ahead state x_{t+2}. Specifically, the RE assumption sets $u_t = E(x_{t+2}|Y_t)$ in (1.1), with Y_t denoting the σ-algebra generated by the output process $\{y_j, j \leq t\}$. A linear *RE* model, where the state variables evolution depends only on past expectations about their future values and on a given random process, is thus obtained as[5]:

$$\tilde{x}_{t+1} = A_t \tilde{x}_t + B_t E(\tilde{x}_{t+2}|\tilde{Y}_t) + v_t, \quad \tilde{x}_0 = \bar{x} \tag{1.3}$$

$$\tilde{y}_t = C_t \tilde{x}_t + w_t, \quad \tilde{y}_0 = \bar{y} \tag{1.4}$$

In De Santis *et al.* (1993) an interesting circumstance, arising under RE, is pointed out. In this situation the "ideal" decision would consist in the *exact* knowledge of x_{t+2}: if the agent were to perfectly foresee the future state values in a world where economic decisions are forecast-dependent, his choice would indeed equal the actual two-step ahead value of the state variables, i.e. $u_t = x_{t+2}$. In this case, we would have an ideal evolution $\{x_t^*\}$ of the state process, governed by the following

[5]Equation (1.3) can be interpreted, for example, as the equilibrium condition of an asset market framework. Note that, given the reliance of the state on past expectations, the latter does not fit into the general setup of Blanchard and Kahn (1980).

second order autoregressive (AR) model[6]:

$$x^*_{t+1} = A_t x^*_t + B_t x^*_{t+2} + v_t \tag{1.5}$$

to which the following (fictitious) measurement equation is attached:

$$y^*_t = C_t x^*_t + w_t \tag{1.6}$$

where it is assumed $x^*_0 = x_0 = \bar{x}$, $x^*_{-1} = 0$.

Since the expectational term at time t depends on the values of the state at future times, the solution of the RE model (1.3)-(1.4) cannot be found by means of substitution of the state values in the state equation. For this reason the model is often called *non-causal*, and the existence of a solution for it is not evident at all (e.g., Broze *et al.*, 1985).

Nevertheless, solutions to (1.3)-(1.4) exist, yet these are in general non-unique[7]. In this work, an algorithm is provided for computing the conditional expectations of the future state $E(\tilde{x}_{t+2}|\tilde{Y}_t)$ which enables us to find a particular solution to a subclass of the general RE models (1.3)-(1.4) - namely those displaying no backward-looking[8] dimension - by estimating the control sequence $\{\hat{u}_t\}$ (among all the admissible Y_t-adapted ones $\{u_t\}$) via a reference model adaptive technique. The *actual* system is indeed forced to "track" a model whose evolution describes the *ideal* situation of perfect foresight of the two-step ahead state value. As the resulting state motion *a fortiori* is the closest - in mean square - to the "ideal" one given by the corresponding perfect foresight model, it can be regarded as the mostly meaningful in the RE philosophy.

As a measure of the closeness to the ideal state motion, it will be used the variance of the difference between the ideal and the actual state at any time, that is:

$$I(u,t) =_{def} E\left((x_t - x^*_t)^T (x_t - x^*_t)\right)$$

[6]Strictly speaking, equation (1.5) belongs to the family of discrete-time descriptor systems, frequently used to model economic processes. For non-singular B_t, the system at issue admits an autoregressive representation.

[7]Basar (1989) offers an example of a scalar and stationary version of the RE model dealt with in the present work, which is satisfied by two different (not stochastically equivalent) random sequences starting from the same initial condition.

[8]That is, in which no lagged values of the state vector enter the right-hand side of the RE equation (1.3).

where E denotes the expectation operator. The dependence of the index $I(u,t)$ on the input sequence $u \equiv \{u_j\}$, $j = 0, 1, \ldots$ comes implicitly from the true state sequence $\{x_t\}$ given by (1.1)[9]. Consistently with the RE hypothesis[10], in our setting the economic agent has full knowledge of the time-varying parameters as well as of the model structure - i.e., of the linear shape of the system and measurement equations (1.1)-(1.2). For the purpose of the analysis, we also require that the model structure be fully captured by the latter, since unmodeled dynamics might seriously worsens the filter performance by causing the estimation algorithm to diverge.

In the next section, a closed loop solution, that is a sequence $\hat{u} = \{u_j\}$ where $\hat{u}_j = \hat{u}(j, y_j, \ldots, y_0)$, is found to the minimization problem of the index $I(u,t)$ for any $t = 0, 1, \ldots$.

1.3 The general solution for the dynamic stochastic model

1.3.1 The "nearly ideal" solution

Let us define the deviation from the ideal state motion x_t^* as:

$$\epsilon_t = x_t - x_t^* \tag{1.7}$$

and consider for a non-negative integer N the following quadratic index:

$$\bar{I}(u,N) = \sum_{t=0}^{N+1} E(\epsilon_t^T \epsilon_t) \tag{1.8}$$

It is straightforward to derive a difference equation for the error ϵ_t using equations

[9] Note that, due to the causality of the model (1.1), this dependence involves only the input values up to time $t-1$. General assumptions such as controllability and observability are required. In particular, no further constraints are considered for the admissible set of instruments.

[10] Under the RE hypothesis, the full state of the economy at all dates and the structure that generated that state are part of the information set upon which expectations are built.

(1.1) and (1.5):

$$\begin{aligned}\epsilon_{t+1} &= x_{t+1} - x^*_{t+1} \\ &= A_t\epsilon_t + B_t\left(u_t - x^*_{t+2}\right) \\ &= A_t\epsilon_t + B_t u_t - x^*_{t+1} + A_t x^*_t + v_t\end{aligned} \tag{1.9}$$

with initial condition $\epsilon_0 = 0$. Let us define the vectors:

$$z_t =_{def} \begin{pmatrix} \epsilon_t \\ x^*_t \\ x^*_{t+1} \\ y^*_t \end{pmatrix}; \quad \zeta_t =_{def} \begin{pmatrix} v_t \\ w_{t+1} \end{pmatrix}$$

so that we can write the following recursive equation for the augmented state z_t, $t \in [0, N+1]$:

$$z_{t+1} = \bar{A}_t z_t + B_{1,t}\zeta_t + B_{2,t}u_t, \quad z_0 = \bar{z} \tag{1.10}$$

where the matrices $\bar{A}_t \in \Re^{3n+m \times 3n+m}$, $B_{1,t} \in \Re^{3n+m \times n+m}$ and $B_{2,t} \in \Re^{3n+m \times n}$ have a block structure with all block-entries being square matrices of dimensions $n \times n$ (see Appendix A.1).

In order to obtain an output equation for (1.10), let us consider equation (1.2) which gives the actual sequence of observations:

$$y_t = C_t x_t + w_t = C_t\epsilon_t + y^*_t = \bar{C}_t z_t \tag{1.11}$$

where $\bar{C}_t \in \Re^{m \times 3n+m}$ has the following block structure:

$$\bar{C}_t = \begin{pmatrix} C_t & 0 & 0 & I \end{pmatrix}$$

Equation (1.11) links the actual observations y_t to the state variables z_t given by (1.10); hence it can be correctly used as the output measurements of the augmented linear discrete-time system in which the first n entries of the state vector describe the evolution of the deviation from the ideal behavior of the RE model. The following result holds:

Property 1. *The new state system with free error measurement, constituted by equations (1.10)-(1.11), has Gaussian state noise.*

Proof. - See Appendix B.1. □

We can solve the problem of finding the set of controls $\{\hat{u}_t\}_{t=0}^N$ which minimizes the objective functional (1.8), as a linear-quadratic-Gaussian (LQG) stochastic control problem on the discrete and finite interval $[0, N+1]$ for the system (1.10)-(1.11). The problem writes then as follows[11]:

$$\min_{\{u_t\}_{t=0}^N} \sum_{t=0}^{N+1} E(z_t^T M z_t) \tag{1.12}$$

subject to the linear system equations (1.10) and observations (1.11), with M being the symmetric matrix with the identity matrix I as first block on the main diagonal and 0's elsewhere. The equations yielding the sequence of optimal controls $\{\hat{u}_t\}$ are in the well known closed-loop form (Bertsekas, 1976):

$$\hat{u}_t = L_t \hat{z}_t \tag{1.13}$$

where $\hat{z}_t =_{def} E(z_t|Y_t)$, Y_t denoting the σ-algebra generated by the observations given by equation (1.11), which are available at time t. The matrices $L_t \in \Re^{n\times 3n+m}$ have the following expression for $t \in [0, N]$:

$$L_t = -(B_{2,t}^T K_{t+1} B_{2,t})^{-1} B_{2,t}^T K_{t+1} \bar{A}_t$$

with K_t given by the following backward recursive equation:

$$\begin{aligned} K_{N+1} &= M \\ K_t &= \bar{A}_t^T \left[K_{t+1} - K_{t+1} B_{2,t} \left(B_{2,t}^T K_{t+1} B_{2,t} \right)^{-1} B_{2,t}^T K_{t+1} \right] \bar{A}_t + M \end{aligned} \tag{1.14}$$

In order to compute the input sequence (1.13), the derivation of the conditional expectations $\hat{z}_t$ is needed. This can be accomplished by means of the Kalman filter equations implemented on the system (1.10)-(1.11), provided $(\bar{A}_t, B_{1,t})$ is a controllable pair, whereas $(\bar{C}_t, \bar{A}_t)$ is an observable one. For any *deterministic* input u_t, the

[11]No penalty (possibly time-varying) weights on the deviations of the state variables from their desired path - the second order AR process (1.5) - are considered, with no loss of generality.

Kalman filter equations, in a form which accounts for the absence of measurement noise (Kalman, 1960; Kailath, 1980), are the following:

$$\hat{z}_{t+1} = \left[I - \bar{K}_{t+1}\bar{C}_t\right]\left[\bar{A}_t\hat{z}_t + B_{2,t}u_t\right] + \bar{K}_{t+1}y_{t+1}$$

where $\bar{K}_t$ is the precomputable filter gain recursively given by[12]:

$$\begin{aligned} \bar{K}_t &= P_{t|t-1}\bar{C}_t^T\left(\bar{C}_t P_{t|t-1}\bar{C}_t^T\right)^{-1} \\ P_{t+1|t} &= \bar{A}_t P_t \bar{A}_t^T + \bar{Q}_t \\ P_t &= P_{t|t-1} - \bar{K}_t\bar{C}_t P_{t|t-1} \end{aligned} \tag{1.15}$$

P_t and $P_{t|t-1}$ being the filtering and one-step prediction error covariances respectively (see Appendix D.1)[13]. As the input gets *stochastic*, and measurable with respect to the observations $(y_1 \ldots y_t)$ up to time t (as in (1.13)), substituting simply $u_t = \hat{u}_t$ gives the optimal filter equations for the case at issue (Carravetta *et al.*, 2002; Liptser and Shiryaev, 2004)[14]:

$$\hat{z}_{t+1} = \left[I - \bar{K}_{t+1}\bar{C}_t\right]\left[\bar{A}_t + B_{2,t}L_t\right]\hat{z}_t + \bar{K}_{t+1}y_{t+1} \tag{1.16}$$

Equation (1.14) for $t = N$ is as follows:

$$K_N = \bar{A}_N^T\left[M - MB_{2,N}\left(B_{2,N}^T M B_{2,N}\right)^{-1} B_{2,N}^T M\right]\bar{A}_N + M$$

and since:

$$\left(B_{2,N}^T M B_{2,N}\right)^{-1} = B_N^{-1} B_N^{-1^T}$$

$$MB_{2,N}\left(B_{2,N}^T M B_{2,N}\right)^{-1} B_{2,N}^T M = M$$

[12]The matrix $\bar{C}_t P_{t|t-1}\bar{C}_t^T$ whose inverse appears in equation (1.15) is always non-singular (see Appendix C.1).

[13]To initialize the prediction error covariance, namely $P_{0|-1} = E(z_0 z_0^T)$, it is sufficient to use the initial condition for the the states z_t.

[14]What is shown in Carravetta *et al.* (2002), is that such property is a fairly 'universal' one in filtering theory, in that it holds for a large class of non-linear filters and non-linear 'feedback' systems (including of course the linear case). However it has being formerly known for a long time to hold in the conditionally-Gaussian case (as reported in Liptser and Shiryaev's textbook). The system at issue lies indeed within the latter.

it follows that $K_N = M$. Therefore we have:

$$K_t = M, \quad 0 \leq t \leq N$$

and the matrices L_t of the controls (1.13) are given by the following formula:

$$L_t = -\left(B_{2,t}^T M B_{2,t}\right)^{-1} B_{2,t}^T M \bar{A}_t \tag{1.17}$$

Insofar as this formula does not depend on the finite horizon N, it yields the optimal control law for all the LQG control problems under the quadratic index (1.12) for any $N = 0, 1, \ldots$. As a control sequence minimizing (1.8) does minimize all the variances $E\left(\epsilon_t^T \epsilon_t\right)$ for any $t = 0, 1, \ldots$, we state the following:

Proposition 1. *Given the model (1.1) with noisy observations (1.2), the sequence of agent's decisions $\hat{u} = \{\hat{u}_t\}$ which produces for any $t = 0, 1, \ldots$ the mean square minimum deviation from the ideal state motion is given by the following equation:*

$$\hat{u}_t = -\left(B_{2,t}^T M B_{2,t}\right)^{-1} B_{2,t}^T M \bar{A}_t \hat{z}_t \tag{1.18}$$

where $\hat{z}_t$ is given recursively by equations (1.15) and (1.16).

We can add further theoretical insight into the nearly ideal control law (1.18), by expressing it as a function of the conditional expectations of the future ideal state, $\hat{z}_t = E(z_t|Y_t)$. To this end, let us make some manipulations on the matrix L_t giving the linear feedback rule for the optimal control:

$$L_t = -\left(B_{2,t}^T M B_{2,t}\right)^{-1} B_{2,t}^T M \bar{A}_t = \begin{pmatrix} -B_t^{-1} A_t & -B_t^{-1} A_t & B_t & 0 \end{pmatrix}$$

which, along with the definition of z_t, can be used in (1.13) to yield:

$$\begin{aligned} \hat{u}_t &= \begin{pmatrix} -B_t^{-1} A_t & -B_t^{-1} A_t & B_t & 0 \end{pmatrix} \hat{z}_t \\ &= -B_t^{-1} A_t E(\epsilon_t|Y_t) - B_t^{-1} A_t E(x_t^*|Y_t) + B_t^{-1} E(x_{t+1}^*|Y_t) \\ &= -B_t^{-1} A_t \hat{\epsilon}_t - B_t^{-1} A_t \hat{x}_t^* + B_t^{-1} \hat{x}_{t+1|t}^* \end{aligned} \tag{1.19}$$

Equation (1.19) allows us to express the optimal control sequence $\{\hat{u}_t\}$ in a more convenient form than (1.18), as stated in the following:

Proposition 2. *For the dynamic stochastic model (1.1)-(1.2), the sequence of agent's decisions* $\{\hat{u}_t\}$ *which produces for any* $t = 0, 1, \ldots$ *the mean square minimum deviation from the ideal solution is given by the following equation:*

$$\hat{u}_t = -B_t^{-1} A_t \hat{z}_{1,t} - B_t^{-1} A_t \hat{z}_{2,t} + B_t^{-1} \hat{z}_{3,t} \tag{1.20}$$

where $\hat{z}_{1,t}, \hat{z}_{2,t}, \hat{z}_{3,t} \in \Re^n$ *are the first three subvectors of* $\hat{z}_t \in \Re^{3n+m}$ *recursively given by equations (1.15) and (1.16).*

The so obtained sequence given by equation (1.20) therefore represents the control rule which produces an evolution of the state x_t featuring the minimum variance displacement ϵ_t from the ideal state x_t^* for any $t = 0, 1, \ldots$. An agent acting in an economic environment as described by equations (1.1)-(1.2) and holding expectations on the future values of the state x_t, will then be sure - by taking the sequence $\{\hat{u}_t\}$ of actions - of exerting the "best" (from his point of view) influence on the future evolution of the economic variables, as this will be the closest one - in mean square - to the ideal evolution the system would have if the agent could act as an "oracle".

1.3.2 The solution for the optimal estimate of the ideal state

Let us now compute the solution of (1.1)-(1.2) when the optimal estimate of the ideal two-step ahead value of the state, i.e. x_{t+2}^*, is used as input. To this end, we first demonstrate the following:

Lemma 1. *Given the linear system (1.1)-(1.2), the* σ*-algebra* $Y_t = \sigma(y_1, y_2, \ldots, y_t)$ *is invariant with respect to all admissible - i.e.,* Y_t*-adapted - control laws* $u = \{u_t\}$.

Proof. - See Appendix E.1. □

Let us now exploit the properties of the augmented system (1.10)-(1.11):

$$z_{t+1} = \bar{A}_t z_t + B_{1,t} \zeta_t + B_{2,t} \hat{x}_{t+2|t}^*, \quad z_0 = \bar{z} \tag{1.21}$$

where, using the definition of z_t and the assertion of Lemma 1[15]:

$$\hat{x}^*_{t+2|t} =_{def} E(x^*_{t+2}|Y_t) = B'E(z_{t+1}|Y_t) = B'\hat{z}_{t+1|t}, \quad B' = [0 \quad 0 \quad I \quad 0]$$

and hence:

$$z_{t+1} = \bar{A}_t z_t + B_{1,t}\zeta_t + B_{2,t}B'\hat{z}_{t+1|t}, \quad z_0 = \bar{z} \tag{1.22}$$

Taking expectations conditional on Y_t yields[16]:

$$\hat{z}_{t+1|t} = \bar{A}_t\hat{z}_t + B_{2,t}B'\hat{z}_{t+1|t} = \left(I - B_{2,t}B'\right)^{-1}\bar{A}_t\hat{z}_t \tag{1.23}$$

In order to obtain the filtering estimate $\hat{z}_t$, we apply the Kalman filter to the state equation (1.22) together with the measurement relation (1.11). Equations (1.22) and (1.11) can be viewed as a linear model driven by the stochastic term $\hat{z}_{t+1|t}$, which is a measurable function of the observations $(y_1 \dots y_t)$ up to time t and hence can be treated as a deterministic one when applying the Kalman filter equations (Carravetta *et al.*, 2002). By using the input sequence $u_t = B'\hat{z}_{t+2|t}$ we obtain, through equation (1.23), the optimal filtering estimate as:

$$\hat{z}_{t+1} = \left[I - \bar{K}_{t+1}\bar{C}_t\right]\left(I - B_{2,t}B'\right)^{-1}\bar{A}_t\hat{z}_t + \bar{K}_{t+1}y_{t+1} \tag{1.24}$$

where the gain $\bar{K}_t$ is given by the same equations (1.15) derived for the nearly ideal case.

Finally, equations (1.23) and (1.24) allows us to recover the one-step prediction optimal minimum variance estimate $\hat{z}_{t+1|t}$ for all t, and exploiting the definition of the vector z_t we can easily extract the optimal prediction estimate of the two-step ahead ideal state $\hat{x}^*_{t+2|t}$.

1.4 The solution of the dynamic RE model

In this section we fully characterize the subset in the general class of RE models (1.3)-(1.4) for which a nearly ideal equilibrium exists. To this end it is sufficient to

[15] The latter enables us to neglect the notation difference between the σ-algebras generated by different (admissible) control functions for the causal model (1.1)-(1.2).

[16] The matrix $(I - B_{2,t}B')$ is always non-singular, as it can be readily verified by direct substitution.

pin down an expression for the conditional expectation $E(\tilde{x}_{t+2}|\tilde{Y}_t)$ which appears in (1.3).

Let us consider the system (1.1)-(1.2) with $u_t = \hat{u}_t$, where $\hat{u}_t$ is the agent's action producing the nearly ideal solution given by Proposition 1. We write the state equation of the model (1.1) at the step $t+i+1$ and for the particular input sequence (1.18):

$$x_{t+i+1} = A_{t+i}x_{t+i} + B_{t+i}\hat{u}_{t+i} + v_{t+i}$$

for any t, $i = 0, 1 \ldots$, and letting Y_t denote the σ-algebra generated by the corresponding observations $\{y_j, j \leq t\}$. Taking conditional expectations yields:

$$\hat{x}_{t+i+1|t} = A_{t+i}\hat{x}_{t+i|t} + B_{t+i}E\left(\hat{u}_{t+i}|Y_t\right) \tag{1.25}$$

since $E(v_{t+i}|Y_t) = 0$ for $i = 0, 1, \ldots$. It is straightforward[17] to see that, given the results obtained in the previous Section, we have:

$$E\left(\hat{u}_{t+i}|Y_t\right) = -B_{t+i}^{-1}A_{t+i}\hat{x}_{t+i|t} + B_{t+i}^{-1}\hat{x}^*_{t+i+1|t} \tag{1.26}$$

Thus, substituting (1.26) into (1.25) we obtain the relation:

$$\hat{x}_{t+i+1|t} = \hat{x}^*_{t+i+1|t}, \quad i = 0, 1, \ldots \tag{1.27}$$

that is, for any t and the input $\hat{u}_t$ the optimal prediction of any future *ideal* state, given the *actual* measurement $(y_0 \ldots y_t)$, is equal to that relative to the *actual* state. For $i = 0$ one obtains:

$$\hat{u}_t = -B_t^{-1}A_t\hat{x}_t + B_t^{-1}\hat{x}_{t+1|t} \tag{1.28}$$

Now, consider the solution of the dynamic stochastic model (2.1)-(2.2) when forcing the model with the sequence $u_t = E\left(x^*_{t+2}|Y'_t\right)$ where Y'_t denotes the σ-algebra generated by the observations $(y'_1 \ldots y'_t)$, which in turn are produced by the model itself up to the current time. Let us denote this solution by $\{x'_t\}$[18]:

$$x'_{t+1} = A_tx'_t + B_tE\left(x^*_{t+2}|Y'_t\right) + v_t, \quad x'_0 = \bar{x} \tag{1.29}$$

[17] See Appendix F.1.

[18] The corresponding observation equation is $y'_t = C_tx'_t + w_t$, with initial condition $y'_0 = \bar{y}$.

Taking expectations conditional on Y_t' we have:

$$E\left(x_{t+1}'|Y_t'\right) = A_t E\left(x_t'|Y_t'\right) + B_t E\left(x_{t+2}^*|Y_t'\right)$$

and hence, since the filtrations satisfy $Y_t = Y_t'$ from Lemma 1:

$$E\left(x_{t+2}^*|Y_t\right) = -B_t^{-1} A_t E\left(x_t'|Y_t\right) + B_t^{-1} E\left(x_{t+1}'|Y_t\right)$$

so that the expression for $E\left(x_{t+2}^*|Y_t\right)$ is in the same form of (1.28), which provides $\hat{u}_t$ as a function of the state and the observations.

Although the noises and the initial conditions in (1.1)-(1.2) under the input sequence $\hat{u}_t$ are the same of those in (1.29) and its corresponding observation equation, the two solutions need not coincide. Indeed, the following is shown:

Lemma 2. *The two state motions generated by the dynamic stochastic model with noisy measurement (1.1)-(1.2), when forced by the control sequences (1.20) and $\hat{x}_{t+2|t}^*$ respectively, differ by a Y-measurable quantity, that is $x_t = x_t' + \xi_{t-1}$ for all t, ξ_{t-1} being a Y_{t-1}-adapted process.*

Proof. - See Appendix G.1. □

The following characterization result is thus derived:

Theorem 1. *In the general class of RE models (1.3)-(1.4), the absence of backward-looking dimension ($A_t = 0$) is necessary and sufficient for the existence of a solution - recursively computable via the causal model (1.1)-(1.2) with initial conditions knowledge only - coinciding both with the nearly ideal solution $x = \{x_t\}, y = \{y_t\}$ and the approximate solution $x' = \{x_t'\}, y' = \{y_t'\}$.*

Proof. - See Appendix H.1. □

This result is summarized in the following:

Proposition 3. *The linear stochastic forward-looking RE model with noisy observations and time-varying parameters:*

$$x_{t+1} = B_t E(x_{t+2}|Y_t) + v_t, \quad x_0 = \bar{x} \tag{1.30}$$

$$y_t = C_t x_t + w_t, \quad y_0 = \bar{y} \tag{1.31}$$

always admits a solution having the property of being the closest (in the mean square sense) to the ideal evolution $\{x_t^*\}$ *given by the corresponding first-order autoregressive process (*B_t *non-singular):*

$$x_{t+1}^* = B_t x_{t+2}^* + v_t, \quad x_0^* = \bar{x}$$

This solution is equal to the ones of the causal dynamic stochastic model (1.1)-(1.2) obtained with the following choices for the control u_t*:*

(i) u_t *is set to the optimal control law with respect to the performance index (1.8);*

(ii) u_t *is set to* $E(x_{t+2}^*|Y_t)$*, where* $Y_t = \{y_j, j \leq t\}$ *is generated by (1.2).*

Several RE models fit into the subclass (1.30)-(1.31), for instance any model in which the t-dated state vector depends upon $t+1$-dated expected states and an exogenous random (zero-mean white Gaussian) process, when the conditional mathematical expectation is formed with respect to the σ-algebra Y_{t-1} (Pesaran, 1989). For this class of RE models, the nearly ideal solution has been shown to be an equilibrium, which is also given by the *causal* model obtained by forcing (1.1) with a suitable control, as derived in the previous sections. This was implicitly conjectured in De Santis *et al.* (1993) for the time-invariant version of (1.3)-(1.4); while such conjecture fails to obtain as a general property of dynamic RE models, Theorem 1 fully characterizes the subset in (1.3)-(1.4) for which it proves true.

1.5 The stationary case and steady state behavior

We now consider a dynamic model as (1.1)-(1.2) with time-invariant parameters:

$$x_{t+1} = Ax_t + Bu_t + v_t, \quad x_0 = \bar{x} \tag{1.32}$$

$$y_t = Cx_t + w_t \quad y_0 = \bar{y} \tag{1.33}$$

where the matrices $A, B \in \Re^{n\times n}$, $C \in \Re^{m\times n}$, and A is stable - i.e., all its roots lie inside the unit circle[19]. The sequences $\{v_t\}, \{w_t\}$ are stationary Gaussian zero-mean white noises mutually uncorrelated, uncorrelated with the initial state $\bar{x}$ and having

[19]Again, we require that B be invertible. Note that $A = 0$ is stable.

covariances $Q \in \Re^{n \times n}$ and $W \in \Re^{m \times m}$ respectively. The initial state $\bar{x}$ has mean zero and known covariance matrix $E(\bar{x}\bar{x}^T) = P_0$. The reference model, giving the ideal solution, satisfies the AR process:

$$x^*_{t+1} = Ax^*_t + Bx^*_{t+2} + v_t, \quad x^*_0 = x_0 \tag{1.34}$$

to which the fictitious measurements $y^*_t = Cx^*_t + w_t$, with $y^*_0 = \bar{y}$, is attached.

The optimal control sequence $\{\hat{u}_t\}$ is then obtained as a particular case of Proposition 2:

$$\hat{u}_t = -B^{-1}A\hat{z}_{1,t} - B^{-1}A\hat{z}_{2,t} + B^{-1}\hat{z}_{3,t} \tag{1.35}$$

where $\hat{z}_{1,t}, \hat{z}_{2,t}, \hat{z}_{3,t} \in \Re^n$ are the first three subvectors of $\hat{z}_t \in \Re^{3n+m}$ and $\hat{z}_t = E(z_t|Y_t)$, Y_t being the σ-algebra generated by $\{y_j, j \leq t\}$. The conditional expectations on the augmented state z_t, needed for the computation of (1.35), are given by the recursive equations:

$$\begin{aligned} \hat{z}_{t+1} = &\left(I - \bar{K}_{t+1}\bar{C}\right)\bar{A}\hat{z}_t + \\ &+ D_1\hat{z}_t + \bar{K}_{t+1}\left(y_{t+1} + D_2\hat{z}_t\right), \quad \hat{z}_0 = E(z_0) \end{aligned} \tag{1.36}$$

where $\bar{K}$ is the precomputable filter gain:

$$\begin{aligned} \bar{K}_t &= P_{t|t-1}\bar{C}^T\left(\bar{C}P_{t|t-1}\bar{C}^T\right)^{-1} \\ P_{t+1|t} &= \bar{A}P_t\bar{A}^T + \bar{Q} \\ P_t &= P_{t|t-1} - \bar{K}_t\bar{C}P_{t|t-1} \\ P(0|-1) &= E(z_0 z_0^T) \end{aligned} \tag{1.37}$$

($D_1 \in \Re^{3n+m \times 3n+m}$ and $D_2 \in \Re^{m \times 3n+m}$ are given in Appendix I.1).

Equation (1.36) results from the application of the Kalman filter to the following system describing the evolution of z_t:

$$z_{t+1} = \bar{A}z_t + B_1\zeta_t + B_2u_t, \quad z_0 = \bar{z} \tag{1.38}$$

$$y_t = \bar{C}z_t \tag{1.39}$$

where the matrices $\bar{A} \in \Re^{3n+m \times 3n+m}$, $B_1 \in \Re^{3n+m \times n+m}$ and $B_2 \in \Re^{3n+m \times n}$ are the time-invariant versions of those presented in Appendix A.1.

It is well known that the stability of the matrix $\bar{A}$ is sufficient to guarantee that the gain $\bar{K}_t$ given by equations (1.37) converges to a finite value as the time index goes to infinity, and it is thereby feasible to exploit the asymptotic approximation of the Kalman filter. In order to check the stability of $\bar{A}$, let us note that it displays the following block-triangular structure:

$$\bar{A} = \begin{pmatrix} A & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{pmatrix}$$

so that, being A stable by hypothesis, we need to check only the block:

$$\bar{A}_{22} = \begin{pmatrix} 0 & I & 0 \\ -B^{-1}A & B^{-1} & 0 \\ 0 & C & 0 \end{pmatrix} = \begin{pmatrix} \bar{A}'_{11} & 0 \\ \bar{A}'_{12} & 0 \end{pmatrix}$$

whose stability is then given by the eigenvalues of:

$$\bar{A}'_{11} = \begin{pmatrix} 0 & I \\ -B^{-1}A & B^{-1} \end{pmatrix}$$

The characteristic equation results in the following:

$$\left| \lambda^2 B - \lambda I + A \right| = 0 \tag{1.40}$$

and hence, if all the $2n$ solutions of (1.40) are within the unit circle of the complex plane, there exists the asymptotic approximation of (1.36) and the steady state gain $\bar{K}$ can be precomputed by iterating (1.37) until convergence is achieved, or equivalently solving the algebraic Riccati equation.

1.6 An extension to Markov-switching RE models

A number of recent studies have made important progress toward connecting the reduced form econometric literature on regime switching autoregressive processes, which can be traced back to Hamilton (1989), with structural economic theory, by developing the notion of Markov-switching Rational Expectations (MSRE) models, that is dynamic forward-looking stochastic frameworks in which the parameters governing the behavior of the system are taken to be functions of a discrete-state Markov chain.

Since able to account for parameter instability and yield quantitatively different responses of macroeconomic variables to fundamental shocks from those implied by fixed regime models, MSRE systems have recently been advocated to investigate the role of regime switching monetary policy in New-Keynesian frameworks (e.g., Davig and Leeper, 2007) or rather to gauge the effects of uncertainty over structural parameters governing the optimal behavior of rational agents (e.g., Liu *et al.*, 2009).

From a technical viewpoint, regime dependency engenders structural nonlinearities which prevent from employing standard solution tools for linear RE systems, such as Blanchard and Kahn (1980)'s, King and Watson (2002)'s and Sims (2002)'s. In this respect, a number of authors have been interested in deriving determinacy (local uniqueness) conditions for RE equilibria to MSRE models. In their seminal contribution to the generalization of the Taylor principle, Davig and Leeper (2007) study how regime switching alters determinacy properties of RE solution and provide analytical restrictions on monetary policy behavior to ensure local uniqueness of the equilibrium path. By focusing on bounded solutions, Davig and Leeper (2007) find out that, while accounting for structural shifts noticeably enlarges the determinacy region relative to the constant parameter setup, regimes that fail to fulfill the generalized Taylor principle may well be characterized by improved time series properties as reaction to fundamental shocks, even when sunspot noise or nonfundamental uncertainty are ruled out. The nonlinearity problem is addressed by introducing a two-step solution method that consists in studying an augmented system which is linear in fictitious variables, the latter coinciding with the actual ones in some of the regimes, and then using the solution to the linear representation in order to construct solutions for the original nonlinear system.

From a more general perspective, Farmer *et al.* (2008, 2009) have provided a series of characterization results for the set of minimal state variable (MSV) solutions as well as the full set of RE equilibria - also sunspot ones - to MSRE frameworks, which satisfy a suitable stability concept. Their approach rests on expanding the state-space of the underlying stochastic system and to focus on an equivalent model in the expanded space that features state-invariant parameters. Furthermore, Farmer *et al.* (2009) demonstrate an equivalence property between determinacy for MSRE models and mean-square stability in a class of Markov jump autoregressive systems.

Here we show how to generalize the Kalman filter approach presented in the previous sections to solution of MSRE models[20]. To this end, let us introduce the following class of purely forward-looking MSRE models with noisy observations on the state vector, defined on a properly filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathcal{P})$:

$$x_{t+1} = \Gamma_{s(t)}^{-1} E[x_{t+2}|\mathcal{F}_t] + \Gamma_{s(t)}^{-1} \Psi_{s(t)} v_t, \quad x_0 = \bar{x} \tag{1.41}$$

$$y_t = \Phi_{s(t)} x_t + w_t \tag{1.42}$$

where x_t is an n-dimensional real vector of random variables of economic interest, y_t is an m-dimensional real vector of observables, and the state error $v_t \in \Re^n$, the measurement noise $w_t \in \Re^m$ and the initial state $\bar{x} \in \Re^n$ are zero-mean white Gaussian processes. With no loss of generality, the covariances of the unobserved structural disturbance and of the measurement noise are normalized to the $I_{n\times n}$ and $I_{l\times l}$ identity matrices respectively, whereas $\bar{x}$ has covariance P_0. $\Gamma_{s(t)}$,$\Psi_{s(t)}$ and $\Phi_{s(t)}$ are conformable matrices holding the coefficients of the underlying economic model, with $\Gamma_{s(t)}$ assumed invertible, as in Farmer *et al.* (2009).

In (1.41)-(1.42), the regime switches are governed by an ergodic discrete-state Markov chain indexed by $s(t)$, with $s(t) \in \mathbb{S} := \{1, \ldots, S\}$. Let $\widetilde{\mathcal{S}}$ denote the σ-field of all subsets in $\mathbb{S}$, and $\widetilde{\mathcal{F}}_t$ the σ-field of $\Re^{n+l}$ in which (x_t, y_t) lie. We define:

$$\Omega := \prod_{t\in\mathcal{T}} \left(\Re_t^{n+l} \times \mathcal{S}_t\right)$$

where $\Re_t^{n+l}, \mathcal{S}_t$ are copies of $\Re^{n+l}, \mathcal{S}$ and $\mathcal{T}$ denotes a discrete-time set of interest. Let $\mathcal{T}_t := \{k \in \mathcal{T}; k \leq t\}$ for each $t \in \mathcal{T}$, then:

$$\mathcal{F} := \sigma\left\{\prod_{t\in\mathcal{T}} (\alpha_t \times \beta_t); \alpha_t \in \widetilde{\mathcal{F}}_t, \beta_t \in \widetilde{\mathcal{S}}, \forall t \in \mathcal{T}\right\}$$

[20]For a generalization of this approach to nonlinear models in which the conditional expectations term is a nontrivial function of the current states and the fundamental shocks, see Sorge (2010a).

and for each $t \in \mathcal{T}$:

$$\mathcal{F}_t := \sigma \left\{ \prod_{\iota \in \mathcal{T}_t} \alpha_\iota \times \beta_\iota \times \prod_{\tau \in \mathcal{T} \setminus \mathcal{T}_t} \Re_\tau^{n+l} \times \mathcal{S}_\tau ; \alpha_\iota \in \widetilde{\mathcal{F}}_\iota, \beta_\iota \in \widetilde{\mathcal{S}}, \iota \in \mathcal{T}_t \right\}$$

with $\mathcal{F}_t \subset \mathcal{F}$. Then $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathcal{P})$ defines a stochastic basis for (1.41)-(1.42), with $\mathcal{P}$ representing a probability measure such that:

$$\mathcal{P}\{s(t+1) = j | \mathcal{F}_t\} = \mathcal{P}\{s(t+1) = j | s(t)\} = p_{s(t)j}$$

with $p_{i,j} \geq 0$ for $i, j \in \mathbb{S}$ and $\sum_{j=1}^{S} p_{ij} = 1$ for each $i \in \mathbb{S}$. The initial conditions $(\bar{x}, s_0)$ are taken to be independent random variables.

More specifically, the information set available at time t, upon which conditional (rational) expectations $E[\cdot|\mathcal{F}_t]$ in (1.41) are built, includes the complete filtrations generated by the output process (1.42), namely $\{y_k, k \leq t\}$, and by the Markov state realizations $\{s_k, k \leq t\}$[21]. We thus allow for observable shifts in modes solely, as in most of the macroeconomic literature on regime switching RE models (e.g., Davig and Leeper, 2007; Farmer *et al.*, 2009)[22]. Accordingly, while the current values of parameters are known, future ones are uncertain. As a working assumption, we also require that $\bar{x}$, v_t, w_t and s_t be mutually independent.

Let the representative agent behave according a specific forecasting function u_t, which need not coincide with RE, and consider the following Markov jump (controllable) system with linear noise corrupted observations:

$$x_{t+1} = \Gamma_{s(t)}^{-1} u_t + \Gamma_{s(t)}^{-1} \Psi_{s(t)} v_t, \quad x_0 = \bar{x} \tag{1.43}$$

$$y_t = \Phi_{s(t)} x_t + w_t \tag{1.44}$$

where u_t is is $\mathcal{F}_t$-measurable, and define the perfect-foresight (Markov jump autoregressive) dynamics where the two-step ahead values of the x_t variables are perfectly anticipated and no (endogenous) forecasting errors are made:

$$x_{t+1}^* = \Gamma_{s(t)}^{-1} x_{t+2}^* + \Gamma_{s(t)}^{-1} \Psi_{s(t)} v_t, \quad x_0^* = \bar{x}, \, x_{-1}^* = 0 \tag{1.45}$$

[21]Note that we assume simultaneous determination of expectations and state shocks v_t, although the latter are taken to be hidden variables.

[22]For theoretical work dealing with unobserved current regimes, see, among others, Andolfatto and Gomme (2003), Leeper and Zha (2003) and Davig (2004).

$$y_t^* = \Phi_{s(t)} x_t^* + w_t \tag{1.46}$$

where both (1.43)-(1.44) and (1.45)-(1.46) are defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathcal{P})$.

Let us introduce:

$$\epsilon_t := x_t - x_t^*, \qquad z_t' := \begin{pmatrix} \epsilon_t' \; x_t^{*'} \; x_{t+1}^{*'} \end{pmatrix}$$

and consider the problem of finding an input sequence $u = \{u_t\}_{t \in \mathbf{T}}$, $\mathbf{T} = [0, T] \subset \mathbb{N}$, $u_t \in U_t$ - with U_t denoting the space of all square-integrable $\mathcal{F}_t$-measurable random vectors - which minimizes the objective functional:

$$J(u) = E \sum_{t=0}^{T+1} (z_t' M z_t) \tag{1.47}$$

under the following state-space recursive constraints:

$$z_{t+1} = A_{s(t)} z_t + B_{s(t)} u_t + C_{s(t)} v_t, \quad z_0 = \bar{z} \tag{1.48}$$

$$y_t = \bar{\Phi}_{s(t)} z_t + w_t \tag{1.49}$$

where M consists of the identity matrix $I_{n \times n}$ as first block on the main diagonal and 0's elsewhere. The expression (1.49) can be properly used as the observation equation for the augmented Markov jump system (1.48) in which the first n entries of the state vector z_t describe the evolution of the deviation from the autoregressive behavior of the MSRE model.

The design of an input sequence $\{\hat{u}_t\}$, $t \in \mathbf{T}$ minimizing (1.47) subject to (1.48)-(1.49) is accomplished by employing an optimal Markov jump feedback controller in conjunction with the minimum mean-square estimate (MMSE) obtained by a time-varying Kalman filter. We indeed show that a separation principle holds for the system at issue - i.e., the optimal input sequence depends on the observed state only through the optimal estimate of the latter. In the classical literature on Markov jump linear quadratic (MJLQ) problems (e.g., Costa *et al.*, 2005), it has been shown that the solution of such problems engenders a twofold set of coupled Riccati equations, each associated to the filtering and control programs respectively. Since these backward-recursive equations cannot be represented as a single higher-dimensional Riccati equation, structural concepts and algorithms from the classical linear theory are not directly applicable to Markov jump systems. While further requirements are generally needed to determine the existence of a steady-state solution for the coupled Riccati equations (e.g., Blair and Sworder, 1975; Chizeck *et al., 1986*; Abou-Kandil *et al.*, 1995), we prove that, when applied to the solution method for MSRE models we propose in this work, this issue vanishes for the Riccati gain admits a simple time-invariant and state-independent representation, both in finite and infinite horizon problems. The following statement clarifies this insight:

Theorem 2. *Given the system (1.43)-(1.44), the input sequence* $\hat{u}_t := \{\hat{u}_t\}$ *which produces for any* $t = 0, 1, \ldots$ *the mean-square minimum deviation from the Markov jump autoregressive state motion (1.45), is in the form:*

$$\hat{u}_t = \Gamma_{s(t)} \hat{x}^*_{t+1|t} \tag{1.50}$$

where the optimal estimate $\hat{x}^*_{t+1|t} := (0\ 0\ I\ 0)' E[z_t|\mathcal{F}_t]$ *is obtained recursively via a time-varying Kalman filter.*

Proof. - See Appendix J.1. □

The estimator of the one-step ahead perfect-foresight state, $\hat{x}^*_{t+1|t}$, is mean-square optimal with respect to the σ-algebra generated by the *actual* measurement process (1.42), the only available data. Our final claim rests on showing that, for any t and the input (1.50), the optimal two-step ahead prediction of the perfect-foresight state x^*_t following the regime switching law of motion (1.45), given the measurement $(y_0, \ldots, y_t)$ and the filtration $\sigma(s^t) = (s_0, \ldots, s_t)$, is equal to that relative to the actual state x_t in (1.43):

Corollary 1. *Let* $x = (x_t), y = (y_t)$ *be the solution of (1.43)-(1.44) under the control law* $\hat{u}_t$*. Then, for any* t *and Markov state* $s(t) = i \in \{1, 2, \ldots, S\}$*, it holds:*

$$\hat{x}_{t+2|t} = \hat{x}^*_{t+2|t} \tag{1.51}$$

Proof. - It readily follows from Theorem 2 and the independence assumption between v_t and s_t. □

Consider now the perfect-foresight Markov jump state motion (1.45). It is easily verified that:

$$\Gamma^{-1}_{s(t)} E[x^*_{t+2}|\mathcal{F}_t] = \Gamma^{-1}_{s(t)} \hat{u}_t$$

which shows, in conjunction with the assertion of Corollary 1, that the optimal feedback controller (1.50) has the same structure of the conditional (rational) expectation term $E[x_{t+2}|\mathcal{F}_t]$, and hence the solution $x = (x_t), y = (y_t)$ of (1.43)-(1.44) with $\hat{u}_t \equiv E[x_{t+2}|\mathcal{F}_t]$ is an REE for the Markov-switching model (1.41)-(1.42). In other words, both in finite and infinite horizon formulations, there always exists an REE $x = (x_t)$ which is computable via a *causal* Markov jump (controllable) system of the form (1.43)-(1.44), when the forcing forecasting function used by the representative agent is designed as the optimal Markov jump feedback control law $\hat{u}$.

In principle, it is desirable to rule out explosiveness in economic behavior by requiring RE equilibria to be stationary. How can this be accomplished in the presence of MSRE models? As emphasized in Farmer *et al.* (2008, 2009), answering this question is not a trivial task. In the following, we establish a simple easy-to-check condition for the mean-square stability of the RE solution as derived in the previous Section. To this end, let $u_t = \hat{u}_t$ in (1.43)-(1.44), and consider the evolution equation for the RE equilibrium:

$$x_{t+1} = \hat{x}^*_{t+1|t} + \Gamma^{-1}_{s(t)}\Psi_{s(t)}v_t \tag{1.52}$$

with:

$$\hat{x}^*_{t+1|t} = \Gamma_{s(t-1)}\hat{x}^*_{t|t-1} + \bar{K}_{t-1}\bar{\Phi}_{s(t-1)}\eta_{t-1} + \bar{K}_{t-1}w_{t-1} \tag{1.53}$$

where $\bar{K}_t$ is the precomputable filter gain and $\eta := z_t - \hat{z}_t$ denotes the estimation error[23].

We study the first two moments of the equilibrium process x_t, i.e. $m_t = E[x_t]$ and $\gamma_t = E[x_t x_t']$, which characterize its mean-square stability. Indeed, system (1.52) is mean-square stable if its first and second moments converge to finite (possibly zero) values in the limit for $t \to \infty$. From (1.52) we have $m_t \to 0$ if and only if $m^*_t = E[\hat{x}^*_{t+1|t}] \to 0$. Moreover, provided that the noise covariance is uniformly bounded with respect to t, i.e. there exists $L \in \Re$ such that:

$$\sum_{i=1}^{S} \|\Gamma_i^{-1}\Psi_i\Psi_i'\Gamma_i'^{-1}\|\mathcal{P}\{s(t) = i\} \leq L < +\infty, \quad \forall t \tag{1.54}$$

[23]See the optimal filter derivation in the proof of Theorem 2.

then $\gamma_t \to 0$ if and only if $\gamma_t^* = E[\hat{x}^*_{t+1|t}\hat{x}^{*'}_{t+1|t}] \to 0$.

Taking expectations in (1.53) yields:

$$E\left[\hat{x}^*_{t+1|t}\right] = E\left[\Gamma_{s(t-1)}\right] E\left[\hat{x}^*_{t|t-1}\right]$$

from which $m_t^* \to 0$ obtains if:

$$\max_i \max_j \left|\lambda_j(\Gamma_i)\right| < 1 \tag{1.55}$$

where $\lambda_j(\Xi)$ denotes the j-th eigenvalue of a matrix Ξ.

As to the second moment, since η_t is orthogonal to $\hat{x}^*_{t+1|t}$ and the measurement noise w_t proves independent of x_t and the σ-algebra $\{y_k, k \le t\}$, we readily derive:

$$\begin{aligned}\gamma_t^* = E\left[\Gamma_{s(t-1)}\gamma^*_{t-1}\Gamma'_{s(t-1)}\right] + E\left[\bar{K}_{t-1}\bar{K}'_{t-1}\right] \\ +E\left[\bar{K}_{t-1}\bar{\Phi}_{s(t-1)}\bar{P}_{t-1}\bar{\Phi}'_{s(t-1)}\bar{K}'_{t-1}\right]\end{aligned}$$

where $\bar{P}_t = E\left[P_t\right]$ and $P_t := E\left[(z_t - \hat{z}_t)(z_t - \hat{z}_t)'|s_k, k \le t\right]$ is the mean-squared error covariance. Thus, $\gamma_t^* \to 0$ for $t \to \infty$ obtains if:

$$\max_i \max_j \left|\lambda_j(\Gamma_i\Gamma_i')\right| < 1 \tag{1.56}$$

and:

$$\sum_{i=1}^{S} \|\Phi_i\Phi_i'\|\mathcal{P}\{s(t) = i\} \le L < +\infty, \quad \forall t \tag{1.57}$$

$$\|\bar{P}_t\| \le L < +\infty, \quad \forall t \tag{1.58}$$

In fact, (1.57) is always verified as $\mathcal{P}$ is a probability measure, and from (1.58) it follows that $\bar{K}_t\bar{K}_t'$ is bounded as well[24]. As for $\bar{P}_t$, its evolution is described by the following recursive equation[25]:

$$\begin{aligned}\bar{P}_{t+1} = E[A_{s(t)}\bar{P}_t A'_{s(t)}] + E[C_{s(t)}C'_{s(t)}] \\ -E[A_{s(t)}\bar{P}_t\bar{\Phi}'_{s(t)}\left(I + \bar{\Phi}_{s(t)}\bar{P}_t\bar{\Phi}'_{s(t)}\right)^{\dagger}\bar{\Phi}_{s(t)}\bar{P}_t A'_{s(t)}]\end{aligned}$$

[24]Note that $E[A_{s(t)}A'_{s(t)}]$ has the same eigenvalues as $E[\Gamma_{s(t)}\Gamma'_{s(t)}]$ and in addition zero eigenvalues.

[25]Given a matrix Ξ, we denote its pseudoinverse by $\Xi^{\dagger}$.

where $\bar{P}_0 = cov(z_0, z_0)$. Riccati equations with Markov jump coefficients have been extensively studied in the engineering literature (e.g., Chizeck *et al.*, 1986; Chizeck and Ji, 1988; Abou-Kandil *et al.*, 1995; Costa *et al., 2005*). According to well-known results established in the mentioned references, condition (1.55) entails condition (1.58). It follows that requirement (1.55) - which also implies (1.56) - is sufficient for the mean-square stability of the obtained RE equilibrium.

1.7 Conclusion

Under the RE hypothesis, subjective forecasts of decision makers are replaced with the mathematical conditional expectation of some future model's equilibrium state with respect to all the information available as to the economic environment. Since rational expectations need to be model-consistent and are endogenously determined, the hypothesis provides little insight into the mechanism of optimal forecasting.

In the present work, we introduce a dynamic policy optimization formulation for forward-looking models, consistent with the RE philosophy, and address the problem of finding a solution to RE models containing past expectations of future states. The general stochastic linear Gaussian vector case with time-varying parameters is considered and a "nearly ideal" solution for a well-specified subset in the class of general RE models (1.3)-(1.4) is accordingly found. In such class, though the dynamic system is not causal, it always admits a solution which can be generated by means of a *causal* model, as conjectured in De Santis *et al.* (1993). Our method thus pins down a unique solution by construction and can be easily applied in estimation. As the equilibrium state motion minimizes a variance index in the set of all the Y_t-adapted input sequences $\{u_t\}$,*a fortiori* it is the mean square closest to the ideal evolution given by the corresponding perfect foresight model, among all the other possible solutions.

De Santis *et al.* (1993) only addresses the problem of state filtering when assuming the dynamic system be driven by the optimal prediction of the ideal state. We show that the optimal prediction of the ideal state does not generate the nearly ideal state motion; nevertheless, we provide a recursive algorithm for computing and filtering the latter, establishing the existence of well-defined link to optimal control issues. Since the so obtained solution represents the *best tracking* of the ideal evolution, our analysis has indeed implications for decision theory under uncertainty, as economic agents can use equation (1.18) in order to take the best forward-looking decision at any time. To assume *rationally formed* expectations means indeed that economic agents attempt to predict the future as well as possible, given all the available information. If not as a normative tool for decision theory, our model can thus be reasonably thought of as a positive description of reality. On the other hand, also from an econometric perspective, equation (1.24) can be exploited in order to estimate the state variables when these are not directly observable, not even at the current time.

Appendix to Chapter 1

A.1 The initial condition for equation (1.10) is given by:

$$z_0 = \begin{pmatrix} \epsilon_0 \\ x_0^* \\ x_1^* \\ y_0^* \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{x} \\ B_0^{-1}\bar{x} \\ C_0\bar{x} + w_0 \end{pmatrix}$$

The matrices $\bar{A}_t \in \Re^{3n+m\times 3n+m}$, $B_{1,t} \in \Re^{3n+m\times n+m}$ and $B_{2,t} \in \Re^{3n+m\times n}$ appearing in the same equation have the following block structure:

$$\bar{A}_t = \begin{pmatrix} A_t & A_t & -I & 0 \\ 0 & 0 & I & 0 \\ 0 & -B_t^{-1}A_t & B_t^{-1} & 0 \\ 0 & 0 & C_{t+1} & 0 \end{pmatrix};$$

$$B_{1,t} = \begin{pmatrix} I & 0 \\ 0 & 0 \\ -B_t^{-1} & 0 \\ 0 & I \end{pmatrix}; \quad B_{2,t} = \begin{pmatrix} B_t \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

B.1 Proof of Property 1 We show that the system (1.10)-(1.11) has Gaussian state noise. Indeed, being the original noise sequences $\{v_t\}$ and $\{w_t\}$ mutually uncorrelated and Gaussian, it follows that they are also jointly Gaussian and $\xi_t = [v_t \, w_{t+1}]^T$ has covariance matrix:

$$E(\xi_t \xi_t^T) = \begin{pmatrix} Q_t & 0 \\ o & W_{t+1} \end{pmatrix}$$

and hence the covariance matrix $\bar{Q}_t$ of the white noise term $B_{1,t}\xi_t$ forcing equation

(1.10) is easily obtained as:

$$\bar{Q}_t = B_{1,t} E(\xi_t \xi_t^T) B_{1,t}^T$$

C.1 We prove that the matrix $\bar{C}_t P_{t|t-1} \bar{C}_t^T$ whose inverse appears in equation (1.15) is always non-singular. Using the second expression in (1.15) we have indeed:

$$\bar{C}_t P_{t|t-1} \bar{C}_t^T = \bar{C}_t \bar{A}_{t-1} P_{t-1} \bar{A}_{t-1}^T \bar{C}_t^T + \bar{C}_t \bar{Q}_{t-1} \bar{C}_t^T$$

and from the definition of the matrices $\bar{C}_t$ and $\bar{Q}_t$ we obtain:

$$\bar{C}_t \bar{Q}_{t-1} \bar{C}_t^T = C_t Q_{t-1} C_t^T + W_t$$

and hence, the non-singularity of the matrix at issue readily follows from taking into account that:

$$\bar{C}_t \bar{A}_{t-1} P_{t-1} \bar{A}_{t-1}^T \bar{C}_t^T \geq 0$$

$$C_t Q_{t-1} C_t^T \geq 0; \quad W_t \geq 0; \quad W_t = W_t^T$$

D.1 In (1.15), P_t and $P_{t|t-1}$ are the filtering and one-step prediction error covariances respectively:

$$P_t = E\left((z_t - \hat{z}_t)(z_t - \hat{z}_t)^T \right)$$

$$P_{t|t-1} = E\left(\left(z_t - \hat{z}_{t|t-1}\right) \left(z_t - \hat{z}_{t|t-1}\right)^T \right)$$

and $\hat{z}_{t|t-1} =_{def} E(z_t | Y_{t-1})$.

E.1 Proof of Lemma 1 We prove that, given the linear system (1.1)-(1.2), the σ-algebra $Y_t = \sigma(y_1, y_2, \ldots, y_t)$ is invariant with respect to all admissible - i.e., Y_t-adapted - control laws $u = (u_t)$. To this end, let us consider for the sake of simplicity the time-independent version of the model (1.1)-(1.2):

$$x_{t+1} = Ax_t + Bu_t + v_t$$

$$y_t = Cx_t + w_t$$

where $x_0 = \bar{x}$ is fixed. Let $u = (u_t)$ and $\tilde{u} = (\tilde{u}_t)$ be two admissible control processes with corresponding solutions $x = (x_t)$, $y = (y_t)$ and $\tilde{x} = (\tilde{x}_t)$, $\tilde{y} = (\tilde{y}_t)$, respectively, and filtrations (Y_t) and $(\tilde{Y}_t)$. For $t = 0$, $y_0 = C\bar{x} + w_0 = \tilde{y}_0$, so that trivially $Y_0 = \tilde{Y}_0$. Proceeding by induction, assume that $Y_s = \tilde{Y}_s$, $s \leq t$, then:

$$y_{t+1} - \tilde{y}_{t+1} = C(x_{t+1} - \tilde{x}_{t+1}) = CA(x_t - \tilde{x}_t) + CB(u_t - \tilde{u}_t) \tag{1.59}$$

But:

$$x_t = A^t \bar{x} + \sum_{i=0}^{t-1} A^i B u_{t-1-i} + \sum_{i=0}^{t-1} A^i v_{t-1-i}$$

and similarly for $\tilde{x}_t$. Therefore:

$$x_t - \tilde{x}_t = \sum_{i=0}^{t-1} A^i B(u_{t-1-i} - \tilde{u}_{t-1-i})$$

By induction hypothesis, $u_s - \tilde{u}_s$ is both Y_s- and $\tilde{Y}_s$-measurable for $s \leq t$. Accordingly, the right-hand side of (1.59) can be written as $g(y_0, \ldots, y_t)$ as well as $\tilde{g}(\tilde{y}_0, \ldots, \tilde{y}_t)$ such that:

$$y_{t+1} = \tilde{y}_{t+1} + \tilde{g}(\tilde{y}_0, \ldots, \tilde{y}_t)$$

and:

$$\tilde{y}_{t+1} = y_{t+1} - g(y_0, \ldots, y_t)$$

This shows the assertion. The same result holds plainly for the time-varying model (1.1)-(1.2) in the main text.

F.1 To obtain equation (1.26), it is sufficient to note that the optimal decision $\hat{u}_t$ has indeed an alternative expression, which easily follows by substituting (1.7) in (1.19), namely:

$$\hat{u}_t = -B_t^{-1} \left[A_t \hat{x}_t - \hat{x}^*_{t+1|t} \right]$$

The result follows exploiting the following identities (which hold to within stochastic equivalence):

$$E\left(\hat{x}_{t+i}|Y_t\right) = E\left(E\left(x_{t+i}|Y_{t+i}\right)|Y_t\right) = E\left(x_{t+i}|Y_t\right) = \hat{x}_{t+i|t}$$

$$E\left(\hat{x}^*_{t+i+1}|Y_t\right) = E\left(E\left(x^*_{t+i+1}|Y_{t+i+1}\right)|Y_t\right) = E\left(x^*_{t+i+1}|Y_t\right) = \hat{x}^*_{t+i+1|t}$$

G.1 Proof of Lemma 2 We show that the two state motions generated by the dynamic stochastic difference model with noisy measurement (1.1)-(1.2), when forced by the input sequence (1.20) and the optimal minimum variance (prediction) estimate of the two-step ahead state, respectively, differ by a Y-measurable quantity, that is $x_t = x'_t + \xi_{t-1}$ for all t, ξ_{t-1} being a Y_{t-1}-adapted process. To this end, let us consider the two systems, given by forcing (1.1)-(1.2) with $\hat{u}_t$ and $u_t = \hat{x}^*_{t+2|t}$, respectively. As said, the same σ-algebra is generated by the corresponding observations equations. According to equation (1.28) we have:

$$\hat{u}_t = -B_t^{-1}\left[A_t\hat{x}_t - \hat{x}_{t+1|t}\right]$$

and:

$$\hat{x}^*_{t+2|t} = -B_t^{-1}\left[A_t\hat{x}'_t - \hat{x}'_{t+1|t}\right]$$

The systems can be rewritten as:

$$x_{t+1} = A_t x_t - A_t E(x_t|Y_t) + E(x_{t+1}|Y_t) + v_t, \quad x_0 = \bar{x} \tag{1.60}$$

$$y_t = C_t x_t + w_t, \quad y_0 = \bar{y} \tag{1.61}$$

and:

$$x'_{t+1} = A_t x'_t - A_t E(x'_t|Y_t) + E(x'_{t+1}|Y_t) + v_t, \quad x'_0 = \bar{x} \tag{1.62}$$

$$y'_t = C_t x'_t + w_t, \quad y'_0 = \bar{y} \tag{1.63}$$

Define the innovations:

$$\tilde{x}_t = x_t - E(x_t|Y_{t-1}), \quad (\tilde{x}_0 = \bar{x})$$

$$\tilde{x'}_t = x'_t - E(x'_t|Y_{t-1}), \quad (\tilde{x'}_0 = \bar{x})$$

so that, from (1.60)-(1.63) one obtains:

$$\tilde{x}_{t+1} = A_t x_t - A_t E(\tilde{x}_t|Y_t) - A_t E(x_t|Y_{t-1}) + v_t, \quad x_0 = \bar{x} \tag{1.64}$$

$$y_t = C_t x_t + w_t, \quad y_0 = \bar{y} \tag{1.65}$$

and:

$$\tilde{x'}_{t+1} = A_t x'_t - A_t E(\tilde{x'}_t|Y_t) - A_t E(x'_t|Y_{t-1}) + v_t, \quad x'_0 = \bar{x} \tag{1.66}$$

$$y'_t = C_t x'_t + w_t, \quad y'_0 = \bar{y} \tag{1.67}$$

Let us assume - the induction hypothesis at time t, which clearly holds at $t = 0$ - that with probability one:

$$x_t = x'_t + \xi_{t-1}$$

with ξ_{t-1} being a Y_{t-1}-adapted process. Then:

$$\tilde{x}_t = x_t - E(x_t|Y_{t-1}) = x'_t - E(x'_t|Y_{t-1}) = \tilde{x'}_t$$

and, by comparing (1.64) against (1.66):

$$\tilde{x}_{t+1} = \tilde{x'}_{t+1}$$

or:

$$x_{t+1} - x'_{t+1} = E(x_{t+1} - x'_{t+1}|Y_t)$$

Defining $\xi_t = E(x_{t+1} - x'_{t+1}|Y_t)$, we finally have:

$$x_{t+1} = x'_{t+1} + \xi_t$$

That is, the two state motions differ by a Y-measurable quantity[26].

H.1 Proof of Theorem 1 To show the assertion, we make use of the following:

Property 2. *The two solutions $x = (x_t)$, $y = (y_t)$ and $x' = (x'_t)$, $y' = (y'_t)$, with filtrations (Y_t) and (Y'_t) coincide if and only if $A_t\hat{\epsilon}_t = 0$.*

Proof. SUFFICIENCY: Assume $A_t\hat{\epsilon}_t = A_t(\hat{x}_t - \hat{x}^*_t) = 0$. From the ideal model (1.5):

$$A_t(\hat{x}_t - \hat{x}^*_t) = 0 \quad \Leftrightarrow \quad \hat{x}^*_{t+1|t} = A_t\hat{x}_t + B_t\hat{x}^*_{t+2|t}$$

[26]Coherently, for the observations we have $y_t = y'_t + \eta_{t-1}$, with η_{t-1} being a Y_{t-1}-adapted process. This result is fully consistent with the fact that such observations generate the same σ-algebra.

and rearranging yields:

$$\hat{x}^*_{t+1|t} = A_t\hat{x}_t + B_t\hat{x}^*_{t+2|t} \quad \Leftrightarrow \quad \hat{x}^*_{t+2|t} = B_t^{-1}[\hat{x}^*_{t+1|t} - A_t\hat{x}_t] = \hat{u}_t$$

Since the initial conditions $(\bar{x}, \bar{y})$ and the system/measurement errors are the same in the two dynamic systems generated by the optimal controls $\hat{u}_t$ and the optimal minimum variance estimate $\hat{x}^*_{t+2|t}$, it follows that $x = x' \, \forall t$.
NECESSITY: Assume now $x_t = x'_t \, \forall t$, then from the ideal model (1.5) one has:

$$\hat{x}^*_{t+2|t} = \hat{u}_t = -B_t^{-1}(A_t\hat{x}_t - \hat{x}^*_{t+1|t}) \quad \Leftrightarrow \quad A_t(\hat{x}_t - \hat{x}^*_t) = 0$$

□

Property 3. *The state motion $x_t = x'_t$ is an RE equilibrium for (1.3)-(1.4) if and only if $A_t\hat{\epsilon}_t = 0$.*

Proof. SUFFICIENCY: Let us consider the following systems:

$$x_{t+1} = A_t x_t + B_t\hat{u}_t + v_t \tag{1.68}$$

and

$$x'_{t+1} = A_t x'_t + B_t\hat{x}^*_{t+2|t} + v_t \tag{1.69}$$

Evidently $x_0 = x'_0$ since for this to hold it is necessary and sufficient that:

$$\hat{x}^*_{1|0} = A_t\bar{x} + B_t\hat{x}^*_{2|0}$$

which is always true given the dynamics (evaluated at $t = 0$):

$$x^*_{t+1} = A_t x^*_t + B_t x^*_{t+2} + v_t \tag{1.70}$$

Assume now $x_s = x'_s, s \leq t$. By induction, for $x_{t+1} = x'_{t+1}$ to hold, it is necessary and sufficient - given (1.68) and (1.69) - that:

$$\hat{x}^*_{t+1|t} = A_t\hat{x}_t + B_t\hat{x}^*_{t+2|t}$$

or equivalently - given (1.70):

$$A_t(\hat{x}_t - \hat{x}_t^*) = 0$$

Provided the latter condition is fulfilled, for any t we have:

$$x_t = x'_t, \quad y_t = y'_t$$

and therefore:

$$E\left(x^*_{t+2}|Y'_t\right) = E\left(x^*_{t+2}|Y_t\right) = \hat{u}_t$$

This shows the equivalence between the nearly ideal solution of subsection (1.3.1) and that obtained in subsection (1.3.2) for $u_t = \hat{x}^*_{t+2|t}$. It follows that equation (1.27) holds true also for such control sequence, yielding - for $i = 1$ - $\hat{x}_{t+2|t} = \hat{x}^*_{t+2|t}$ and hence the solution of (1.1)-(1.2) with:

$$u_t = \hat{x}^*_{t+2|t} = \hat{u}_t = E(x_{t+2}|Y_t)$$

is a solution of the RE model (1.3)-(1.4).
NECESSITY: Assume $x_t = x'_t$ is an RE equilibrium, then it must solve:

$$\hat{x}_{t+1|t} = A_t\hat{x}_t + B_t\hat{x}_{t+2|t}$$

Since x_t is generated by the optimal control $\hat{u}_t$, equation (1.27) holds true and the previous RE model proves equivalent to:

$$\hat{x}^*_{t+1|t} = A_t\hat{x}_t + B_t\hat{x}^*_{t+2|t}$$

which, given the dynamics (1.5), implies $A_t\hat{\epsilon}_t = 0$. □

Sufficiency of the claim in Theorem 1 follows directly from Property 2 and 3. To show necessity, it is sufficient to note that, when $x_t = x'_t$, from the recursive equations:

$$x_{t+1} = A_t x_t + B_t L_t \hat{z}_t + v_t$$

$$x'_{t+1} = A_t x'_t + B_t B'(I - B_{2,t}B')^{-1}\bar{A}_t\hat{z}_t + v_t$$

with L_t given by (1.17), it follows necessarily:

$$-(B_t^T B_t)^{-1} B_{2,t}^T \bar{A}_t = B'(I - B_{2,t}B')^{-1}\bar{A}_t$$

which defines a set of parametric restrictions of the form:

$$-(B_t^T B_t)^{-1} B_t^T A_t = 0 \tag{i}$$

$$-(B_t^T B_t)^{-1} B_t^T A_t = -B_t^{-1} A_t \tag{ii}$$

$$(B_t^T B_t)^{-1} B_t^T = B_t^{-1} \tag{iii}$$

Noticing that (ii) and (iii) are identities[27], from (i) it follows that $A_t = 0$.

I.1 The matrices $D_1 \in \Re^{3n+m \times 3n+m}$ and $D_2 \in \Re^{m \times 3n+m}$ in equations (1.37) are given by:

$$D_1 = \begin{pmatrix} -A & -A & I & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad D_2^T = \begin{pmatrix} CA \\ CA \\ -C \\ 0 \end{pmatrix}$$

J.1 Proof of Theorem 2 To save notation, let us define $u_t^+ = [u_t', u_{t+1}', ..., u_T']'$, and let $\xi^t = [\xi_0', ..., \xi_t']'$ denote a sequence of random vectors $\xi_0, ..., \xi_T$. The σ-algebra generated by $\xi_0, ..., \xi_t$, namely $\sigma(\xi^t)$, will be for simplicity identified with the vector ξ^t.

We first derive the conditional expectations for the augmented state vector z_t. This is accomplished by employing a time-varying Kalman filter for the state-space system (1.48)-(1.49). Indeed, the objective is to identify at every time step t, an estimate $\hat{z}_t$ that minimizes the mean-squared error covariance:

$$P_t = E\left[(z_t - \hat{z}_t)(z_t - \hat{z}_t)'|s^t\right]$$

A potential issue lies in that the noise provides information about the state since

[27] Indeed $(B_t^T B_t)^{-1} B_t^T$ is a pseudoinverse (in this case, actually the inverse) of B_t.

the regime switching matrices multiplying the two depend on the same underlying Markov state. However, as long as the current realization of the Markov chain is observable, the state variable z_t and the noise v_t become independent. Likewise, though the noise turns correlated, conditioned on the current state estimate and the Markov state, the next period noise remains (conditionally) zero-mean.
Since the estimator at time t has access to observations $(y_0, ..., y_t)$ and the Markov state values $(s_0, ..., s_t)$, the optimal linear MMSE filtering estimate $E[z_t|\mathcal{F}_t]$ is obtained from a time-varying (sample path) Kalman filter (e.g., Chizeck and Ji, 1988). Let $s_t = i \in \mathbb{S}$ be the Markov state observed in time t, then:

$$\hat{z}_t = \hat{z}_{t|t-1} + \bar{K}_t \left(y_t - \bar{\Phi}_i \hat{z}_{t|t-1} \right), \quad \hat{z}_0 = E\{z_0\} \tag{1.71}$$

$$\bar{K}_t = P_{t|t-1} \bar{\Phi}_i' \left(I + \bar{\Phi}_i P_{t|t-1} \bar{\Phi}_i' \right)^{\dagger} \tag{1.72}$$

$$\hat{z}_{t+1|t} = A_i \hat{z}_t + B_i u_t$$

$$P_t = P_{t|t-1} - \bar{K}_t C_i P_{t|t-1}$$

$$P_{t+1|t} := E\left[(z_{t+1} - \hat{z}_{t+1|t})(z_{t+1} - \hat{z}_{t+1|t})' | s^t \right] = A_i P_t A_i' + C_i C_i'$$

where $P_0 = cov(z_0, z_0|s_0)$.
Using the measurement equation (1.49), (1.71) rewrites:

$$\hat{z}_{t+1} = A_i \hat{z}_t + B_i u_t \bar{K}_t \left(\bar{\Phi}_i (z_t - \hat{z}_t) + w_t \right)$$

which along with (1.48) yields the equation of the estimation error $\eta_t := z_t - \hat{z}_t$:

$$\eta_{t+1} = \left(A_i - \bar{K}_t \bar{\Phi}_i \right) \eta_t + C_i v_t - \bar{K}_t w_t \tag{1.73}$$

from which we observe that η_t is independent of u_t.
We turn now to the Markov jump LQG problem described by (1.47)-(1.48)-(1.49). Let us define the cost-to-go at t:

$$J_t(u_t^+, \mathcal{F}_t) = E\left\{ \sum_{s=t}^{T+1} z_s^T M z_s \middle| \mathcal{F}_t \right\} \tag{1.74}$$

and the optimal cost-to-go (at t):

$$J_t^*(\mathcal{F}_t) = \min_{u \in U} J_t(u_t^+, \mathcal{F}_t) \tag{1.75}$$

where U readily follows from the above defined U_t, and the min is taken samplewise with respect to $\mathcal{F}_t$. Finally denote:

$$u_t^{+*} = \arg\min_{u \in U} J_t(u_t^+, \mathcal{F}_t) \tag{1.76}$$

The optimality principle ensures that $\left(u_t^{+*}\right)_{t+1}^{+} = u_{t+1}^{+}{}^{*}$, i.e. the restriction of the optimal control sequence for the t-th instance of the sequence (1.75) of optimal control problems, is the optimal control for the $t+1$-th problem. Straightforward computation yields the following recursive relation between the optimal cost-to-go functionals (1.75):

$$J_t^*(\mathcal{F}_t) = E\left\{z_t' M z_t | \mathcal{F}_t\right\} + \min_{u_t} E\left\{J_{t+1}^*(\mathcal{F}_{t+1}) \middle| \mathcal{F}_t\right\} \tag{1.77}$$

which is the general equation of the Dynamic Programming Algorithm (DPA). Going backwards, at the last stage one has:

$$u_0^{+*} = \arg\min_{u \in U} J_0(u_0^+, \mathcal{F}_0)$$

hence a fortiori:

$$u_0^{+*} = \arg\min_{u \in U} E\left\{J_0(u_0^+, \mathcal{F}_0)\right\} = \arg\min_{u \in U} J(u)$$

which delivers the desired solution.
As to the initial stage, we need $J_T^*(\mathcal{F}_T)$, which requires us to solve for:

$$u_T^* = \arg\min_{u_T} J_T(u_T, \mathcal{F}_T) = \arg\min_{u_T} E\left\{z_T' M z_T + z_{T+1}' M z_{T+1} \middle| \mathcal{F}_T\right\} \tag{1.78}$$

and then to substitute it into the functional:

$$
\begin{aligned}
& J_T^*(I_T) = J_T(u_T^*, \mathcal{F}_T) \\
& = E\left\{ z_T' M z_T + z_{T+1}' M z_{T+1} \middle| \mathcal{F}_T \right\} \\
& = E\Big\{ z_T' M z_T + z_T' A_{s(T)}' M A_{s(T)} z_T + u_T^{*\prime} B_{s(T)}' M B_{s(T)} u_T^* \\
& \qquad + 2 z_T' A_{s(t)}' M B_{s(T)} u_T^* + v_T' C_{s(T)}' M C_{s(T)} v_T \Big| \mathcal{F}_T \Big\}
\end{aligned}
$$

where it has been used the independence of $z_T, s(T)$ and v_T, which implies:

$$
E\left\{ z_T' A_{s(T)}' M C_{s(T)} v_T | \mathcal{F}_T \right\} = E\left\{ z_T' A_{s(T)}' M C_{s(T)} E\{v_T\} | \mathcal{F}_T \right\} = 0 \tag{1.79}
$$

as well as:

$$
E\left\{ u_T' B_{s(T)}' M C_{s(T)} v_T | \mathcal{F}_T \right\} = E\left\{ u_T' B_{s(T)}' M C_{s(T)} E\{v_T\} | \mathcal{F}_T \right\} = 0 \tag{1.80}
$$

by the independence of s^T, y^T, hence of $u_T \equiv u_T(\mathcal{F}_T)$, and v_T.
Noting that u_T only affects the quadratic form of z_{T+1} in (1.78), thus using the system equation, it holds:

$$
u_T^* = \arg\min_{u_T} E\left\{ z_{T+1}' M z_{T+1} \middle| \mathcal{F}_T \right\} \tag{1.81}
$$

Using (1.79), (1.80), and noting that u_T does not affect the quadratic terms in z_T and v_T, we obtain:

$$
\begin{aligned}
u_T^* \quad & = \arg\min_{u_T} E\left\{ u_T{}' B_{s(T)}' M B_{s(T)} u_T + 2 z_T' A_{s(t)}' M B_{s(T)} u_T \middle| \mathcal{F}_T \right\} \\
& = \arg\min_{u_T} \left\{ u_T{}' B_{s(T)}' M B_{s(T)} u_T + 2 \hat{z}_T' A_{s(t)}' M B_{s(T)} u_T \right\}
\end{aligned}
$$

By setting to zero the derivative respect to u_T of the positive quadratic functional in the above equation, and solving with respect to u_T, we get u_T^*:

$$
u_T^* = -\left(B_{s(T)}' M B_{s(T)} \right)^{-1} B_{s(T)}' M A_{s(T)} \hat{z}_T \tag{1.82}
$$

and substituting (1.82) into (1.78), the following expression of the optimal cost at time T obtains:

$$J_T^*(I_T) = E\left\{ z_T' K_T z_T + (z_T - \hat{z}_T)' L_T (z_T - \hat{z}_T) + v_T' C_{s(T)}' M C_{s(T)} v_T \Big| \mathcal{F}_T \right\} \quad (1.83)$$

where:

$$L_T = A_{s(T)}' M A_{s(T)} \quad (1.84)$$

$$K_T = M - L_T + A_{s(T)}' M A_{s(T)} = M \quad (1.85)$$

Now, the DPA (1.77) for $t = T-1$ implies:

$$\begin{aligned} u_{T-1}^* &= \arg\min_{u_{T-1}} E\left\{ J_T^*(\mathcal{F}_T) \big| \mathcal{F}_{T-1} \right\} \\ &= \arg\min_{u_{T-1}} E\left\{ z_T' K_T z_T \big| \mathcal{F}_{T-1} \right\}, \\ &= \arg\min_{u_{T-1}} E\left\{ z_T' E\{K_T\} z_T \big| \mathcal{F}_{T-1} \right\} \end{aligned}$$

where the second equality comes from being the estimation error $(z_t - \hat{z}_t)$ not affected by u_t, and the third one from being $z_T, \mathcal{F}_{T-1}$ independent of $s(T)$. Hence, a recursive representation arises for the problem at hand, with the following general characterization of the optimal control holding true:

$$u_t^* = \arg\min_{u_t} E\left\{ z_t' E\{K_t\} z_t \big| \mathcal{F}_{t-1} \right\}$$

whose value is given by:

$$u_t^* = -\left(B_{s(t)}' E\{K_t\} B_{s(t)} \right)^{-1} B_{s(t)}' E\{K_t\} A_{s(t)} \hat{z}_t \quad (1.86)$$

where the gain K_t solves the backward-recursive equations:

$$L_t = A_{s(t)}' E\{K_{t+1}\} B_{s(t)} \left(B_{s(t)}' E\{K_{t+1}\} B_{s(t)} \right)^{-1} B_{s(t)}' E\{K_{t+1}\} A_{s(t)} \quad (1.87)$$

$$K_t = E\{K_{t+1}\} - L_t + A_{s(t)}' E\{K_{t+1}\} A_{s(t)}, \quad K_{T+1} = M \quad (1.88)$$

As M is a square, idempotent matrix, from (1.85) it follows that $K_t = M$ for all periods $t = 1, \cdots, T$ and states $s(t) \in S$.

Finally, by substitution of K_t in (1.86) we derive[28]:

$$u_t^* = \Gamma_{s(t)} \hat{x}_{t+1|t}^* \tag{1.89}$$

Insofar as the expression for the feedback matrices does not depend on the finite horizon T, it yields the optimal control law for all the LQG control problems in the (1.47)-(1.48)-(1.49) form for any $T = 1, 2, \ldots$.

[28] The third entry of $\hat{z}_t$ is $E[x_{t+1}^* | \mathcal{F}_t]$.

Chapter 2

On the Identifiability of News Shocks and Sunspot Models under Rational Expectations

2.1 Introduction

In recent years, economists have been paying great attention to the idea that advance information on shifts in fundamentals, as opposed to current changes in opportunities or preferences, may configure an important driver of aggregate fluctuations (e.g., Beaudry and Portier, 2006; Schmitt-Grohé and Uribe, 2008; Jaimovich and Rebelo, 2009). While the classic sunspot literature (e.g., Azariadis, 1981; Cass and Shell, 1983; Benhabib and Farmer, 1994, 1999) emphasizes the importance of forecast revisions driven by extrinsic (non-fundamental) uncertainty in rational expectations (RE) frameworks, these studies rather investigate the possibility of expectations-driven fluctuations in regular economies when accounting for an information structure under which forthcoming developments in the economy are (possibly imperfectly) anticipated.

From a theoretical perspective, it has been demonstrated that the two types of shocks to beliefs generally involve different cross-equation restrictions and implications for the dynamic properties of equilibria, and that models in which sunspots are able to generate business cycle comovements may well fail to do so under news shocks (e.g., Karnizova, 2007). A subsequent, fundamental research step appears then to be the extension of these results to econometric issues, notably to the analysis of whether stochastic systems driven by news shocks can be empirically distinguished from ones which allow for sunspots.

The purpose of the present study is to show, in the same spirit of Beyer and Farmer (2008), that expectations-driven models are plagued by a critical lack of identifiability, since indeterminate equilibrium models other than those featuring news shocks possess the same likelihood function as the latter. The observationally equivalent frameworks are generally characterized by a lagged expectations structure, which can be associated with different microfoundations, like the presence of staggered-price setting under past information (e.g., Woodford, 2003), information stickiness (e.g., Mankiw and Reis, 2002) or imperfect information in monetary policy making (e.g., McCallum and Nelson, 2000). More generally, lagged expectations terms arise in DSGE models whenever (some of the) control variables are predetermined, or based on past information with respect to current observables. By exploiting the general martingale difference solution approach (e.g., Broze and Szafarz, 1991), we show that, for any exactly identified news shocks model with VARMA equilibrium reduced form, an observationally equivalent class of lagged expectations linear RE (LRE) systems exists which is subject solely, though arbitrarily (via indeterminacy), to i.i.d. fundamental shocks.

A key assumption of the literature on DSGE models with news shocks, the one which crucially departs from standard business cycle analysis, is that forward-looking agents are endowed with a richer information set than the one containing current and past realizations of exogenous variables. Indeed, future shocks to the latter are typically assumed to be - possibly only in part - anticipated (e.g., Schmitt-Grohé and Uribe, 2008). Interestingly, the observational equivalence result established in this work involves stochastic difference models featuring some degree of information stickiness, according to which past state variables partly determine the current behavior of endogenous variables, and the largest σ-field upon which conditional forecasts are built - i.e. the one containing current and past observables - is only a subset of the one available to economic agents acting in news shocks frameworks.

The main implication of introducing lagged expectations terms in LRE systems is that serially correlated sunspot variables arise in equilibrium reduced forms. Wang and Wei (2006) were the first to point out this fact. However, Wang and Wei (2006)' solution algorithm for LRE systems with lagged expectations configures a particular case of the general martingale difference approach presented in Broze and Szafarz (1991), as the former restricts attention to linear stationary solutions and thus implicitly excludes arbitrariness in the equilibrium MA coefficients, which can be pinned down from structural parameters knowledge. We rather follow Broze and Szafarz (1991) to exploit the possibility of equilibrium forecast errors that feature an arbitrary correlation structure with respect to fundamental shocks[1].

The analysis presented here is more closely related to the econometric literature on the identification and estimation of indeterminate equilibrium RE systems. Observational equivalence is a general issue in RE models, as noticed in the the early

[1]That is, the analysis presented here does not rely on the existence of non-fundamental sunspot noise, as in Beyer and Farmer (2007) and Sorge (2010b).

work of Sims (1980) and Pesaran (1989) and recently emphasized by Beyer and Farmer (2007, 2008). It lies indeed on the largely untestable nature of identifying restrictions on the random processes which generate exogenous variables or the dynamic structure of RE frameworks. We contribute to this literature by extending the equivalence result to news shocks versus indeterminate equilibrium LRE models. Given the prominent position gained by theories regarding news shocks and anticipation effects in the business cycle debate, this finding points to the opportunity of supplementing likelihood-based empirical investigations of news shocks in estimated DSGE models with testing strategies for the indeterminacy hypothesis.

The remaining chapter is organized as follows. In section 2.2, we illustrate how the macroeconomic literature has introduced news shocks and anticipation in DSGE modeling. Section 2.3 introduces the identification issue by looking at a highly stylized model economy under perfect anticipation. In section 2.4 a general equivalence result for multivariate news shocks LRE models is then established. section 2.5 offers concluding remarks. For the sake of exposition, all the proofs and major technical details are relegated to the Appendix.

2.2 Modeling news shocks in macroeconomics

Following the seminal work of Cochrane (1994), theories of news-driven business cycles have recently gained a prominent position in the macroeconomic debate. Beaudry and Portier (2006), via a structural vector-autoregressive (VAR) approach, has found a non-negligible role for news shocks to total factor productivity (TFP) as a source of economic fluctuations in most of aggregate variables (stock prices, output, consumption, investment and hours)[2]. Several studies have been then developed to assess the importance of advance information for business cycles and explore its implications for policy making (e.g., Beaudry and Portier, 2004, 2007; Schmitt-Grohé and Uribe, 2008; Jaimovich and Rebelo, 2009).

At a purely conjectural level, the anticipation effect is easily understood by

[2]While sensibly improving the time series properties of dynamic stochastic models, news shocks have been shown to introduce non-invertible MA components into equilibrium reduced forms (e.g., Leeper and Walker, 2011). As a consequence, the success of the SVAR approach to proper uncovering this type of shock is controversial (Leeper *et al.*, 2008; Fève *et al.*, 2009).

introducing a simple stochastic process x_t which evolves as:

$$x_t = \rho x_{t-1} + z_t \tag{2.1}$$

with $|\rho| < 1$. Assume this process forces a single-equation RE equation of the form:

$$y_t = aE_t[y_{t+1}] + bx_t \tag{2.2}$$

where y_t is an endogenous variable of economic interest[3]. According to the RE hypothesis, the term $E_t[y_{t+1}] \equiv E[y_{t+1}|\mathcal{F}_t]$ represents the conditional mathematical expectations of y_{t+1} with respect to all the information available at time t and included in the information set $\mathcal{F}_t$, where typically $\mathcal{F}_t = \{a, b, y_{t-j}, x_{t-j}|j = 0, 1, \cdots\}$. However, a core assumption of the mentioned literature on news shocks is that the information set upon which the economic agents build their expectations, is much larger than one simply containing current and past realizations of the disturbance impinging on the exogenous variable. The term z_t is indeed assumed (e.g., Schmitt-Grohé and Uribe, 2008) to be a linear combination of some components which are anticipated several periods in advance as well as a surprise (unanticipated) innovation at time t:

$$z_t = v_{0,t} + v_{t,NEWS} \quad \forall t$$

where, to allow for the variation in the timing of the arrival of the news, the anticipated components are further decomposed as $v_{t,NEWS} = \sum_{j=1}^{q} v_{j,t-j}$, and $v_{j,t-j}$ denotes j-period ahead news on changes in the level of z_t anticipated at period $t-j$, with $q > 1$ being the longest horizon over which anticipation occurs. Accordingly, $v_{j,t-j}$ is in the period $t - j$ information set of economic agents and yet results in an actual change in the level of z_t only in period t.

As to the statistical properties of the news shocks, it is commonly assumed that the disturbances $v_{j,t}$ are i.i.d, normal with zero mean and finite variance σ_j^2, $j \in [1, q]$. This assumption implies zero correlation between the news and contemporaneous shocks as well as zero cross-correlation among news shocks[4]; in other words, $v_{j,t}$ is

[3] Equation (2.2) may be given several economically meaningful interpretations. See, e.g., Blanchard and Fischer (1989).

[4] For a wider representation of news shocks models, which allows for serially correlated news processes, see Leeper and Walker (2011).

uncorrelated across time and across anticipation horizon:

$$E[v_{j,t}v_{i,t-n}] = 0, \qquad i = j \in [1, q], \quad n > 0$$

$$E[v_{j,t}v_{i,t}] = 0, \quad j \neq i$$

The forcing error term z_t is therefore unconditionally mean zero and serially uncorrelated, that is $E[z_t] = 0$ and $E[z_t z_{t-n}] = 0$, $n > 0$; moreover, it is unforecastable given only past realizations of itself, since $E[z_{t+n}|z_t, z_{t-1}, \cdots] = 0$, $n > 0$.

The mentioned approach to modeling news shocks allows for a first-order autoregressive representation for the exogenous process of the form:

$$\hat{x}_t = M\hat{x}_{t-1} + \xi_t \tag{2.3}$$

where[5]:

$$\hat{x}_t = (x_t \quad v_{1,t} \quad v_{2,t} \quad v_{2,t-1} \quad \cdots \quad v_{q,t} \quad \cdots \quad v_{q,t-q+1})'$$

$$\xi_t = (v_{0,t} \quad v_{1,t} \quad v_{2,t} \quad 0 \quad \cdots \quad v_{q,t} \quad 0 \quad \cdots \quad 0)'$$

For example, assuming anticipation occurs up to three periods ahead ($q = 3$), the exogenous process can be written as:

$$\begin{pmatrix} x_t \\ v_{1,t} \\ v_{2,t} \\ v_{2,t-1} \\ v_{3,t} \\ v_{3,t-1} \\ v_{3,t-2} \end{pmatrix} = \begin{pmatrix} \rho & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ v_{1,t-1} \\ v_{2,t-1} \\ v_{2,t-2} \\ v_{3,t-1} \\ v_{3,t-2} \\ v_{3,t-3} \end{pmatrix} + \begin{pmatrix} v_{0,t} \\ v_{1,t} \\ v_{2,t} \\ 0 \\ v_{3,t} \\ 0 \\ 0 \end{pmatrix}$$

This representation illustrates the information updating structure and the propagation mechanism of news shocks through conditional expectations of future values of the exogenous variable x_t. Being forward-looking, at any period t the economic

[5] M is a $(1 + \sum_j^q j) \times (1 + \sum_j^q j)$ square matrix whose first entry is ρ whereas the others are either 1 or 0 in order to recover all the anticipated shocks $v_{j,t-j}$ present in (2.1). The vector of innovations ξ_t is normal i.i.d with zero mean and finite variance-covariance matrix.

agents acting upon the information set:

$$\mathcal{F}_t = \{x_t, z_{1,t}, \dots, z_{j,t}, z_{1,t-1}, \dots, z_{j,t-1}, \dots\}$$

rationally react even before the anticipated shocks are actually realized. Given the best forecast of future values of the error term z_t:

$$E_t[z_{t+j}] = \sum_{i=j}^{q} v_{i,t-i+j} \quad j \leq q$$

the expected value of x_{t+j} at time t can be easily obtained as:

$$E_t[x_{t+j}] = \rho^j x_t + \rho^{j-1} \sum_{i=1}^{j} v_{i,t+1-i} + \rho^{j-2} \sum_{i=2}^{j} v_{i,t+2-i} + \ldots + v_{j,t} \quad j \leq q$$

A slightly different approach to formalizing the endogenous arrival of new information requires integrating filtering algorithms, like the Kalman one or least squares projection, in order to extract signals out of a noisy environment. More precisely, this approach (e.g., Beaudry and Portier, 2004; Andersen and Beier, 2005; Karnizova, 2007), which enables us to specify the anticipated components of the error term forcing the exogenous variable(s) as one-step ahead expectations revisions, models anticipation and information updating within a richer information structure where exogenously given signals are assumed to convey noisy news on future realizations of the error variable z_t; this allows the economic agents to gather some (possibly imperfect) information about incoming innovations to the exogenous states by solving a signal extraction program[6]. In every period t, the agents therefore observe, in addition to current (and past) realizations of the fundamental shock z_t in (2.1), an exogenous (noisy) signal s_t^j as part of an information flow on the $j \in [1, q]$ periods ahead disturbances z_{t+j}:

$$s_t^j = z_{t+j} + u_t, \quad u_t \sim N(0, \sigma_u^2)$$

with u_t being uncorrelated with the corresponding vector of fundamental impulses z_t at all lags and leads.

[6]That is, the agents are endowed with informative signals and learn as an unintended consequence of such observations (passive learning).

While being unable to disentangle the two components of the signals they receive, all the agents collect the same public signal so that the noise term does not vanish in the cross-section average. This non-trivial information problem and consequently the way new information enters the economy is crucial for the behavior and the dynamic adjustment process of macroeconomic systems. The conditional (rational) expectations in period t of future states x_{t+i}, $i \geq 1$ are indeed not entirely determined by the history of their realizations, i.e. it is generally $E_t[z_{t+i}] \neq 0$. Least squares projection implies:

$$E_t[z_{t+i}] \equiv E[z_{t+i}|s_t^i, s_{t-1}^i, ...] = \Theta s_{t+i-j} \quad if \quad i \leq j \tag{2.4}$$

where $\Theta \equiv \sigma_v^2/(\sigma_v^2 + \sigma_u^2)$.

It should be noted that the terms "news" and "noise" are not necessarily restricted to denote (possibly imperfect) information agents receive today as to developments in the economy tomorrow. Part of the literature indeed label "news" as new information which becomes available at the beginning of each time period. For instance, Lorenzoni (2009) and Blanchard *et al.* (2009) build models of demand-side economic fluctuation allowing for noise shocks - in the form of imperfect signals on current, not observable (exogenous) variables - to affect public beliefs. In what follows, we will not be concerned with this latter branch of literature.

2.3 News shocks or correlated sunspots? A simple example

We first consider a simple single-equation model under a news shock, studied in Fève *et al.* (2009). Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathcal{P})$ be a filtered probability space, and y_t be a scalar endogenous variable adapted to $\mathcal{F}_t$ which evolves according to[7]:

$$y_t = \frac{\phi}{1+\phi\gamma} y_{t-1} + \frac{\gamma}{1+\phi\gamma} E_t(y_{t+1}) + \frac{1}{1+\phi\gamma} z_t \tag{2.5}$$

$$\phi \in (-1, 0) \cup (0, 1), \quad |\gamma| < 1$$

where $E_t(\cdot) := E(\cdot|\mathcal{F}_t)$ denotes the conditional expectation operator. The realization

[7] Structural parameters are chosen so as to ensure the existence of a (locally) unique RE equilibrium.

of the exogenous variable z_t is assumed to be fully anticipated one period in advance (news shock):

$$z_t = v_{t-1}, \quad v_t \sim i.i.d.\, N(0, \sigma_v)$$

If the sequence for $E(y_t)$ is bounded[8], the following determinate stationary solution obtains, that only involves fundamentals, as expressed in ARMA (1,1) form:

$$(1 - \phi L)y_t = \gamma(1 + \gamma^{-1}L)v_t \tag{2.6}$$

where L denotes the lag operator. The non-invertibility of the MA component suggests that statistical inference based on simple autoregressions is generally invalid (e.g., Leeper *et al.*, 2008).

Model identification also involves difficulties for empirical analysis. Consider the following RE model featuring a lagged expectation but no news shocks[9]:

$$y_t = a_{01}E_t(y_{t+1}) + a_{11}E_{t-1}(y_t) + bv_t \tag{2.7}$$

$$a_{01} \neq 0, \quad a_{11} \neq 1$$

The reduced form of (2.7) is given by:

$$\sum_{i=0}^{1} a_i^* y_{t+i-1} = \lambda_t - bv_{t-1} \tag{2.8}$$

$$a_0^* = a_{11} - 1, \qquad a_1^* = a_{01}$$

where $\lambda_t = \sum_{k=0}^{1} a_{k1}\xi_{t-k}$, with $\xi_t := y_t - E_{t-1}(y_t)$ being an arbitrary martingale difference sequence with respect to $\mathcal{F}_t$. Thus, model (2.7)'s equilibrium is subject to a correlated sunspot shock λ_t.

Let us assume that $\xi_t = \pi v_t$ for some arbitrary parameter $\pi \in \Re$. The forecast errors can indeed be expressed as $\xi_t = \pi v_t + \delta s_t$, i.e. as a linear combination of the fundamental shock and an extraneous (non-fundamental) sunspot variable s_t satisfying $E_{t-1}(s_t) = 0$, which is assumed to be observed by the agents[10]. The

[8]This is usually enforced by the existence of a transversality condition in the underlying optimization framework.

[9]We will be more specific about the role of restrictions imposed to the alternative model when presenting the general result in the following section.

[10]That is, $\mathcal{S}_t := \sigma(s_k, k \leq t) \subset \mathcal{F}_t$.

impact parameters π and δ are to be selected endogenously; as demonstrated in Lubik and Schorfheide (2003), not all the reduced form parameters are uniquely determined under equilibrium indeterminacy.

Here, we focus on the case where $s_t := 0$ almost surely, $\forall t$ (parametric indeterminacy)[11]. Then from (2.8) we obtain the ARMA (1,1) representation:

$$\left(1+\frac{a_0^*}{a_1^*}L\right)y_t=\pi\left(1+\frac{(a_{11}\pi-b)}{a_1^*\pi}L\right)v_t \tag{2.9}$$

Let $\pi=\gamma$. Then for any element in the set:

$$S:=\left\{(a_{01},a_{11},b):\, a_{01}=\frac{(\gamma-b)}{(1+\phi\gamma)};\, a_{11}=1-\phi a_{01};\, b\neq\gamma\right\}$$

which is non-empty[12], model (2.5) and (2.7) generate the same reduced form forecasts, and thus are observationally equivalent.

It is important to emphasize the role of the lagged expectation term $E_{t-1}(y_t)$ in this simple model economy. Equilibrium indeterminacy allows for self-fulfilling expectations revisions even in the absence of any fundamental shock in (2.7), i.e. when $b=0$ (e.g., Benhabib and Farmer, 1994). Then observational equivalence would obtain for any (non-fundamental) sunspot variable $\xi_t := s_t \sim i.i.d.\, N(0,\gamma\sigma_v)$[13]. However, in this case the presence of a lagged expectation is crucial for the result, as it creates room for correlated sunspots to arise in equilibrium.

2.4 A general equivalence result for multivariate news shocks models

In this section, we discuss the possibility of empirically evaluate the relative importance of different types of beliefs shocks on aggregate macroeconomic variables using the predictions generated by the theory. The main question is whether it is possible for an econometrician looking at time series data $\{Y_t\}$ to decide if the latter

[11] This is the case when the equilibrium reduced form for y_t is to have a moving average structure with respect to the innovation v_t. Indeed, for such a type of solutions the following property holds: $\xi_t^j := E_t(y_{t+j})-E_{t-1}(y_{t+j})=\pi_j v_t, \forall j\in\mathcal{N}$. However, not all the real parameters π_j are necessarily arbitrary (see Broze and Szafarz, 1991).

[12] In particular, the restrictions imposed to (2.7) are preserved.

[13] While it would be harder to justify coordination of agents' expectations revisions on fundamental shocks if the latter were not present, this result is still consistent with the RE framework.

are generated by a model with a determinate equilibrium under news shocks or rather by an indeterminate equilibrium model which allow for correlated sunspots.

As an example, assume that the econometrician were to be confronted with data exhibiting high volatility and persistence, so that he might arguably conjecture that such time series properties shall be ascribed to the presence of anticipated shocks to fundamentals (e.g., Fève *et al.*, 2009). In what follows, we point out that such conjecture may happen to be incorrectly validated by the data when observational equivalence arises between stochastic linear models driven by different types of beliefs shocks. Hence, the theory may fail to provide actually testable identifying restrictions.

For the purpose of our analysis, we assume that the news shocks model is *exactly identified*, i.e. we assume that unique values for the structural parameters can be recovered from the estimates of the reduced form and from independent linear restrictions suggested by the theory (e.g., Rothenberg, 1971; Beyer and Farmer, 2008). Let $Y_t = [y_{1t}, \ldots, y_{nt}]'$ be a vector-valued endogenous variable, and let the following VARMA process define the equilibrium reduced form of a general multivariate RE model featuring (potential) news shocks $Z_t = [z_{1t}, \ldots, z_{nt}]'$, whose j-th element is $z_{jt} = v_{j,t-s_j}$, $\{s_j \in \{0, 1, \ldots, q\}\,; V_t = [v_{1t}, \ldots, v_{nt}]', V_t \sim i.i.d.\, N(0, \Sigma_v)\}$, where q represents the maximum anticipation horizon:

$$(I - \Phi L)Y_t = \Gamma(L)V_t, \qquad \Gamma(L) = \Gamma_0 + \sum_{i=1}^{q} \Gamma_i L^i \tag{2.10}$$

with Γ_0 non-singular. We prove the following[14]:

Theorem 1. *Let $G_0(E_{\mathcal{F}}(Y_-, Y, Y_+), Z; \Theta_0) = 0$ be an exactly identified news shocks RE model with structure $\{G_0, \Theta_0\}$, whose equilibrium reduced form is the VARMA process (2.10). Then a class $G_1(E_{\mathcal{F}}(Y_-, Y, Y_+), V; \Theta_1) = 0$ of observationally equivalent RE models exists, which is only subject to the i.i.d. shocks V_t.*

Proof. - See Appendix A.2. □

Theorem 1 shows that a family of lagged expectations LRE models under i.i.d. fundamental shocks always exists, which generate the same likelihood function as the news shocks model. This identification failure is inherent to the RE hypothesis (e.g.,

[14]The t-dated information set is $\mathcal{F}_t := \sigma(Y_k, k \leq t)$. The time the forecast is formed cannot be posterior to the dating period of the expected variables.

Sims, 1980)[15]. It allows to construct alternative economies driven by different types of expectations shocks which cannot be disentangled empirically. This observational equivalence problem is analogous to that discussed in Beyer and Farmer (2007), in which an econometrician with no independent information on the true variances of the (unobservable) fundamental shocks and the (unobservable) sunspot variables may well be unable to uncover the latter from the real world data.

One way of coping with this problem may consist in testing the hypothesis that the data are generated by an indeterminate equilibrium model. In this respect, testing strategies which are able to control for dynamic misspecification - i.e. the omission of lags, expectational leads or variables with respect to the actual data generating process (Fanelli, 2010) - should be preferred over system-based ones which exploit information on autocovariance patterns of observed time series to deliver evidence on determinacy versus indeterminacy (Lubik and Schorfheide, 2004)[16].

2.5 Conclusion

The main goal of this chapter was to study identification of linear dynamic RE models under news shocks. The main question addressed was whether these models are empirically distinguishable from lagged expectations RE systems which allow for equilibrium correlated sunspots. By means of the general martingale solution approach, it is shown that, for any exactly identified news shocks model, there exists an observationally equivalent class of indeterminate RE systems, which are only subject to i.i.d. fundamental shocks.

Since the alternative models possess different determinacy properties, different implications for policy-making are also likely to arise. We believe that this issue should be carefully addressed in likelihood-based estimation exercises intended to evaluate the quantitative importance of news shocks and anticipation effects as drivers of business cycles.

[15] A simple corollary of this result is that not only may alternative specifications of news shocks in RE models - i.e. i.i.d news and correlated news processes - have identical information content to rational agents (e.g.,,, Leeper and Walker, 2011), but even lead to observationally equivalent equilibrium dynamics.

[16] This latter approach to testing for indeterminacy is in fact affected by dynamic misspecification problems. A different direction is followed by studies on how to address the non-invertibility property of news shocks models (e.g.,, Kriwoluzky, 2009; Dupor and Han, 2011).

Our analysis might also serve as a robustness check for Bayesian studies on indeterminacy testing (e.g., Lubik and Schorfheide, 2004). In fact, enhanced time series properties could still be reproduced with parameters from the determinacy region when news shocks are present, as they are typically associated with endogenous propagation.

Appendix to Chapter 2

A.2 Proof of Theorem 1 The proof works as follows. We first introduce a general multivariate LRE system without news shocks whose reduced form admits the existence of forecast errors in VMA representation, which can be assumed to be arbitrarily related with the fundamental shocks V_t. A set of conditions, in the form of linear restrictions, is then found on this correlation structure and the parameters of the alternative multivariate RE system in order to construct an observationally equivalent model. The assertion is proven by showing that this system of constraints always admits solutions.

To begin with, let us consider the following LRE system:

$$G_1(E_{\mathcal{F}}(Y_-, Y, Y_+), V; \cdot) = 0 :$$

$$Y_t = \sum \sum\nolimits_{(k,i)\in\mathcal{I}} A_{k,i+k} E_{t-k}(Y_{t+i}) + BV_t \tag{2.11}$$

where $A_{k,i+k}$ are conformable matrices and:

$$\mathcal{I} := \{(k, i = h - k) : \, k \in \{0, 1, \ldots, K\} \, , \, h \in \{0, 1, \ldots, H\}\}$$

that is, h and k refer to the lead of the expectations and to the conditioning σ-field $\mathcal{F}_{t-k}$.

Any solution to (2.11) satisfies the recursive system:

$$\sum_{i=J_0}^{J_1} A_i^* Y_{t+i} = \sum_{k=0}^{K} \sum_{h=1}^{H} \sum_{j=0}^{h-1} A_{kh} \Xi_{t+h-k-j}^{j} - BV_t \tag{2.12}$$

where:

(i) $A_0^* = -I + \sum_{k:(k,0)\in\mathcal{W}\rceil} A_{kk}, \quad A_i^* = \sum_{k:(k,i)\mathcal{W}\rceil} A_{k,i+k}, \; i \neq 0;$

(ii) $J_0 = \min \{i \in \mathcal{Z} : \, A_i^* \neq 0\}, \quad J_1 = \max \{i \in \mathcal{Z} : \, A_i^* \neq 0\};$

(iii) $-K \leq J_0 \leq 0 \leq J_1 \leq H;$

(iv) $\Xi_t^j := E_t(Y_{t+j}) - E_{t-1}(Y_{t+j})$ are H n-dimensional (forecasts) revision processes.

However, (2.12) is an equilibrium reduced form of (2.11) if and only if a set of constraints is imposed to the revision processes Ξ_t^j, i.e. the components of the latter may not generally be chosen as arbitrary martingale difference sequences. Let us set $J_1 = H = 1$, $J_0 = 0$ and $K = q$. Equation (2.12) then rewrites:

$$A_1^* Y_t + A_0^* Y_{t-1} = \Lambda_t - BV_{t-1}, \qquad A_i^* \neq 0 \quad i = 0, 1 \tag{2.13}$$

where $\Lambda_t = \sum_{k=0}^{q} A_{k1} \Xi_{t-k}$.This reduced form only involves $(n - n_1)$ arbitrary martingale differences as components of the revision process Ξ_t, where n_1 is the number of zero roots of the following characteristic equation:

$$det \left(\sum_{i=J_0}^{J_1} A_i^* \mu^{J_1 - i} \right) = 0 \tag{2.14}$$

The matrix A_0^* is generically invertible. By requiring A_1^* to be non-singular[17], no zero root exists, and Ξ_t is fully arbitrary. Serially correlated sunspot variables Λ_t then enter model (2.11)'s equilibrium reduced form. Given arbitrariness of forecasts revision processes, we can set $\Xi_t = \Pi V_t$ for some matrix $\Pi \in \Re^{n \times n}$. Then, for any non-singular A_1^* matrix and $\Pi = \Gamma_0$, we can obtain a system of $n^2(2q+1)$ linear restrictions[18] of the form:

$$A_{11} = I - A_{01} \Phi \tag{2.15}$$

$$B = A_{11} \Gamma_0 - A_{01} \Gamma_1 \tag{2.16}$$

$$A_{i1} = A_{01}^{-1} \Gamma_i \Gamma_0^{-1}, \qquad i \in \{2, \ldots, q\} \tag{2.17}$$

$$A_{i-1,0} = -A_{i1}, \qquad i \in \{2, \ldots, q\} \tag{2.18}$$

$$A_{q0} = 0 \tag{2.19}$$

in exactly $n^2(2q+1)$ 'unknowns' (i.e. the structural parameters of the alternative RE model)[19]. Clearly, once A_1^* is chosen, system (2.15)-(2.18) always admits a solution, which we label Θ_1. Along with (2.11), this yields the class of observational equivalent models $G_1(E_{\mathcal{F}}(Y_-, Y, Y_+), V; \Theta_1) = 0$.

[17]This step involves no loss of generality for the elements of A_0^* are really free parameters.

[18]Of which, $n^2(K - J_0) = n^2 q$ are zero-restrictions imposed on some of the A_i^* matrices by the assumption $H = J_1 = 1$ and $J_0 = 0$.

[19]Equivalently, we could have imposed $A_{i-1,0} = A_{i1} = 0$ in (2.18).

Part II

Political Economy

Chapter 3

The Role of the Judiciary in the Public Decision Making Process

3.1 Introduction

A leading concern regarding democratic political systems is the leverage that special interest groups may claim on actual policies by means of political influence-buying and monetary contributions to policy-making institutions. This phenomenon is observed to be pervasive in modern democracies and has gained a prominent position in the political economy debate.

The role of political influence was indeed noted since the middle of last century in the literature on public choice (e.g., Buchanan and Tullock, 1962; Olson, 1965; Hillman and Katz, 1987) and political economy of trade policy and protection (e.g., Hillman, 1982, 1989). The seminal contributions of Stigler (1971), Grossman and Helpman (1994, 2001) and Dixit *et al.* (1997), have provided a characterization of the public decision maker as an auctioneer who may receive bids from various entities, in the form of bribes, campaign contributions, or other alluring incentives[1].

In some political systems, notably the United States, the provision of contributions to politicians may be perfectly legal and considered to be lobbying, whereas in other systems the same transfers would be regarded as illegal and accordingly identified as bribery. Somewhat surprisingly, while in most of the literature lobbying and bribery can be viewed as the same phenomenon[2], little attention has been paid to the role of the judiciary in preventing illegal rent-seeking through bribery.

In fact, while focusing on judicial review as part of the checks and balances against the misuse of political power by the executive branch (e.g., Hanssen, 2004),

[1]The literature on the causes and the consequences of corruption on social welfare is by now a large chapter of public economics, as reviewed in Jain (2001) and Aidt (2003).

[2]The differences between lobbying and bribing have not been extensively addressed in the theoretical literature; in the pioneering work of Grossman and Helpman (1994), lobbying takes the form of monetary transfers from lobbies to politicians, which could equally be interpreted as bribes (e.g., Coate and Morris, 1999). Harstad and Svensson (2011) attempt to draw the boundary by tackling the question why firms choose to lobby - aiming at changing existing rules or policies - or bribe - attempting to get around existing rules or policies -, and the consequences of this choice in a growth framework.

the political economy literature on corruption has generally neglected to investigate the interrelations between the judicial oversight of the policy-making process and the incidence of political corruption. Intuitively, the scope of corruption depends on the expected benefits to politicians from becoming corrupt; hence, systematic and well-targeted efforts toward the investigation and prosecution of bribery cases might serve as a powerful device for corruption deterrence. However, effective oversight of the political process typically faces institutional and operational constraints, like political interference or judicial subversion, which may hinder the proper functioning of this mechanism.

A better understanding of corruption then calls for a careful analysis of the interplay among institutions in shaping the incentives for bribery. While the internal organization of the state affects the strategic behavior of organized groups, it may also alter the incentives for policy-makers to abuse power in their own interest. In principle, the presence of a judiciary that oversees the market for bribes should foster corruption deterrence. However, independent judiciaries that act in corrupt societies are vulnerable to taking bribes (Glaeser and Shleifer, 2002); hence, mechanisms for enforcing judicial accountability are of crucial importance.

This work aims at shedding some light on these issues by developing an endogenous policy framework that captures the interplay among a policy-maker who allocates public funds from a fixed budget between two groups, a judiciary that oversees the political process and investigates corruption charges and a lobby group that may bribe the policy-maker to bias the allocation of funds in its favor and/or the judiciary to exert less effort in investigating corruption in the allocation process. We characterize the political equilibria of the model when accounting for both this form of multiplicity of actors involved in the process of policy-making and the possibility that, while being independent of the political authority, the judiciary itself is bribed by the lobby group[3].

Existing work dealing with corruption in the judicial branch has mostly focused on the corruption of law enforcers and its implications for the deterrence effect of laws (e.g., Becker and Stigler, 1974); on the related issues of optimal monitoring and compensation schemes for law enforcers (e.g., Polinsky and Shavell, 2001) and of

[3]Aiming at influencing the judicial choice in their favor, i.e. toward a less tightening oversight activity to set up. The case in which interest groups face the decision of whether they should lobby the political bodies to switch policy, or rather challenge existing policy before the courts, is developed by Rubin *et al.* (2001).

optimal regulation in the presence of corrupt contract enforcers (e.g., Immordino and Pagano, 2010); and on the general contracting problem under judicial agency from a theoretical perspective (e.g., Bond, 2009). More closely related to our approach is the recent work of Priks (2011), who examines how judicial dependence influences corruption at different levels of the government in a model in which the central authority, the low-level officials and the judiciary are potentially corrupt. Similarly to our model, Priks (2011) argues that even highly corrupt (independent) judiciaries may reduce corruption. The two works, however, differ in several respects. First, we analyze the relationship between judicial independence and political corruption by means of a menu-auction model in which a lobby group acts as the principal of both the policy-maker and the judiciary. While Priks (2011) focuses on explaining how independent judiciaries affect the distribution of rents between officeholders and central authorities, we rather study how the presence of corruptible judiciaries that oversee the political process impacts on the mechanism by which lobby groups can influence policy outcomes, i.e. bribery. Second, our model formalizes some relevant features that are commonly presumed to exert influence on judicial decision-making, like the efficiency of courts. Lastly, Priks (2011) does not deal with the issue of judicial accountability, when independent judges are vulnerable to external influence.

The independence of judiciaries may in fact facilitate corruption in this branch because no other government entity has the authority to oversee them (e.g., Rose-Ackerman, 1978; Glaeser and Shleifer, 2002). We show that judicial independence is a necessary condition for deterrence effects to arise from the oversight activity of judiciaries. In fact, dependent judges are not able to prevent the interest group and the government body from maximizing the profits from the deals between them. This set of results complements those of Landes and Posner (1975), who regard the existence of an independent judiciary as a key element in the successful functioning of political systems where public policies emerge from the attempts of interest groups to influence political decisions in their favor.

While the independence of the judiciary is crucial to its effectiveness, judges must also be held accountable for their institutional role. Our analysis suggests that preserving the efficiency of independent judiciaries can serve as an instrument for self-enforced judicial accountability, even in the presence of corrupt judges. This finding is in line with Dal Bó *et al.* (2006), who point out that well-functioning judicial systems increase the cost of corrupt deals, whereas slow and/or ineffective judicial systems raise the incentives for engaging in corrupt behavior.

We also show that the trade-off between judicial accountability and judicial discretion vanishes when the judiciary is dependent on the political authority. When the policy-maker is excessively concerned with political contributions, entrusting the former with power over the judiciary prevents the existence of (possibly partial) deterrence equilibria. Conversely, the presence of an independent judiciary, even if corrupt, breaks the exclusive bargaining channel with the political authority, and thus weakens the lobby's incentives to engage in bribery - the latter not being able to create large rents. This in turn might reduce the total corruption[4].

The remaining chapter is organized as follows. Section 3.2 reviews empirical literature on the relationship between judicial agency and political corruption. The theoretical model is introduced in section 3.3, while section 3.4 carries out the equilibrium analysis. Section 3.5 illustrates the case of a dependent judiciary. Section 3.6 concludes. For the sake of exposition, all the proofs and major technical details are relegated to Appendix.

3.2 Background

The starting point of our analysis is the question of whether the structure of the judiciary plays a role in determining political corruption, under the assumption that judges themselves may be bribe-takers. This section documents several empirical facts that our framework is able to explain jointly.

Empirical contributions to the study of judicial systems suggest that judicial dependence plays an important role in explaining high levels of corruption. Ades and Di Tella (1997) and La Porta *et al.* (2004) show that political influence over judicial institutions typically increases corruption. More generally, a few strands of econometric research demonstrate the beneficial effects of judicial independence on economic growth and social welfare, developing numerous indicators and providing evidence that countries with strong independence of judicial institutions enjoy higher economic performance and political freedom.

[4]This insight is reminiscent of Rose-Ackerman (1978)'s argument that heightening the number of individuals who must be bribed in order to achieve the desired outcome may in fact be socially preferable.

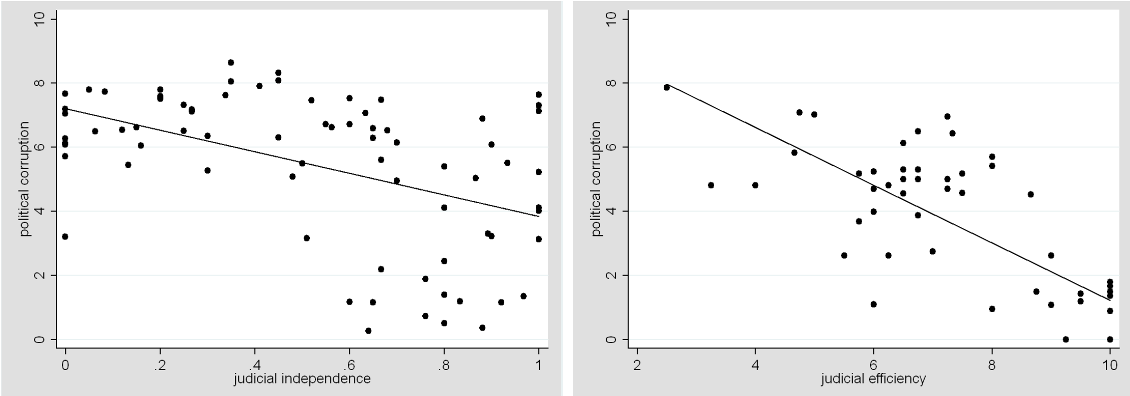

Figure 3.1. Political Corruption and the judiciary.
Source: Aaken *et al.*, 2010 (left panel); Djankov *et al.*, 2002 (right panel). The original indexes of political corruption are rescaled in order that higher scores stand for higher levels of corruption.

The seminal contribution of Feld and Voigt (2003) introduces a twofold notion of judicial independence - *de jure* independence, as described in the constitutional establishment of the supreme court, and *de facto* independence, that is judicial independence as it is actually implemented in practice; exploiting a cross-sectional sample they present evidence that only *de facto* judicial independence is conducive to growth. Recently, Aaken *et al.* (2010) take on an estimation strategy - based on the construction of two ad hoc indicators of independence - in order to test the hypothesis of whether government power over prosecutors may raise government members' incentives to misuse such power in order to prevent the prosecution of illegal activities or crimes - like corruption - committed by themselves. In Figure 1 (left panel) we plot the relationship between (*de facto*) judicial independence and political corruption for a cross-section of countries. The empirical evidence shows that the factual independence of judges is a robustly significant determinant of cross-national variation in political corruption.

Even if the concept of *de facto* independence refers to the factual implementation of judicial independence beyond its formal provision, this feature is not sufficient to cover the overall effectiveness of the judicial role. In particular, the action of the judiciary depends on its quality as well as its autonomy from external influence, in particular from the private sector. There is no shortage of evidence about judicial corruption, though it is often anecdotal in character. Several instances, especially in the U.S., are indeed given in which the latter has been detected and its existence proven ex-post (see Bond, 2009). Substantial systematic evidence points to the fact that the rule of law does not apply in many countries and that judicial decisions are in fact subject to influence (e.g., Boudreaux and Pritchard, 1994).

To our knowledge, the existing theoretical work has tended to focus on each of these features separately. We attempt to bridge this gap by developing a conceptual framework encompassing this larger set of institutions and focusing on two determinants of quality of the judiciary: efficiency and integrity. The efficiency of the judiciary is understood as the ability of a court system to process criminal cases in a professional manner and at a reasonable cost and time. A number of potential determinants have been considered in the literature: adequate endowment of judges and equipment, training of staff, legal formalism and characteristics of the case management system. Even if efficiency has multiple facets, it is nonetheless measurable, unlike some of the other features (Dakolias, 1999)[5].

A second dimension of judicial quality is integrity. If the independence of the judiciary regards autonomy from the political power, the integrity refers to its degree of corruptibility by other special interests. Just as judicial efficiency fosters the rule of law, the integrity of its members guarantees the fairness and impartiality of the prosecution system. According to selected literature (e.g., Mauro, 1995), the relationship between the quality of the judiciary and the presence of political and/or bureaucratic corruption can be regarded as a stylized fact. For example, Figure 1 (right panel) shows linear correlations between an index of judicial efficiency and the level of corruption in a cross-section of countries (obtained by Djankov *et al.*, 2002).

It is important to emphasize that different legal systems provide quite different incentives to judges. In particular, a framework of civil/common law could influence the quality of the judiciary or its dependence from political power and also the degree of corruptibility within the system. Our model formalizes the interaction between the quality and the independence of the judiciary in order to shed some light on their joint effect on the public decision-making process in the presence of illegal rent-seeking activities. Hence, for the sake of simplicity independence, efficiency and integrity of the judiciary are taken to be partly exogenous.

[5]An important contribution to the analysis of courts' behavior is Djankov *et al.* (2002), who collect and investigate data on judicial activity from a large sample of countries.

3.3 The model

We consider a simple economy populated by N individuals divided into two groups of size n_k, with $\sum_{k=1}^{2} n_k = N$. Utility (welfare) is derived from the consumption of pure group-specific public goods. Under homogeneous preferences within each group we have:

$$U_k = n_k G_k(q_k B) \tag{3.1}$$

where $q_k \geq 0$, with $\sum_k q_k = 1$, is the share of a fixed budget B allocated between the groups as group-specific public goods, and $G_k(\cdot)$ is a twice-differentiable function satisfying $G_k'(\cdot) > 0$ and $G_k''(\cdot) < 0$ for $k = 1, 2$. With no loss of generality, we normalize $B = 1$[6].

The effective redistribution scheme results from the interplay of two government institutions. While public policies are unilaterally determined by the decisions of a politician (P), their determination is influenced by the behavior of a separate institution, the judiciary (J), which is given the role of overseeing the political process. In a setting *á la* Grossman and Helpman (2001), an organized interest group k may decide to make its political contribution contingent on the selected policy by formulating a transfer schedule $T_k(q_k)$. The schedule maps any feasible value for the shares $q_k \in [0, 1]$ into a non-negative contribution to P. In our simple economy, contributions are illegal (outright bribes). As P will choose the policy vector $(\hat{q}_1, \hat{q}_2)$ that maximizes its own objective, the joint (net) welfare of the members of the lobby group k is given as

$$V_k = U_k(\hat{q}_k) - T_k(\hat{q}_k) \tag{3.2}$$

We model a reduced form for P's objective function, assuming that fixed weights are exogenously assigned to the welfare levels of the two different groups in the economy[7]. When choosing the tax revenue shares q_k to be allocated for the production of the public goods, the policy-setting authority is thereby concerned with the public's

[6]More generally, we could let $B = t \sum_k n_k y_k$ and $u_k = (1-t)y_k + G_k(q_k B)$, where y_k denotes (exogenously given) gross income and t is the tax rate. None of the results derived herein depend on our simpler modeling choices.

[7]Here, as in the standard literature on endogenous policy, θ_k reflects political relevance and may represent population or electoral weights.

well-being and with the receipts it gets from the groups of interest. We explicitly allow for uncertainty in the payment of the contributions. This mirrors the degree to which the policy-maker is actually captured by the lobby group and is linked to the oversight of the judiciary. This feature is modeled by letting P benefit from the effective transfers only with (known) probability $f \in [0, 1]$. Accordingly, a risk-neutral politician chooses $q \in [0, 1]$ to maximize[8]:

$$V_P = f \sum_{k=1}^{2} T_k(q_k) + l \sum_{k=1}^{2} \theta_k V_k(q_k)$$

where $\sum_k \theta_k V_k(\cdot)$ is the social welfare function with weights $\theta_k > 0$ for $k = 1, 2$ and $\sum_k \theta_k = 1$, while $l > 0$ denotes the (exogenous) degree of preference of P for social welfare relative to bribes.

Groups may differ in their ability to capture institutions and outbid rival seekers of favorable policies. We assume that group 1 only acts as a bribe provider, and let group 2 represent the unorganized general public (e.g., Olson, 1965). To simplify notation, we set $q_1 = q$ (and $q_2 = 1 - q$ accordingly), so that $T(q)$ and $\widehat{T} \equiv T(\hat{q})$ will denote the bribe schedule with which the lobby confronts the politician and the effective transfer eventually paid by the lobby in exchange for the chosen policy $\hat{q}$, respectively. The payoff function of P then reduces to

$$V_P = fT(q) + l \sum_{k=1}^{2} \theta_k V_k(q_k) \tag{3.3}$$

The judiciary is in charge of an anti-corruption office which is incidental to the effective transfer of the bribe $\widehat{T}$. The incentives of the judiciary are shaped by its internal structure as well as by the institutional interplay with the political body. Following Posner (1994, 1995), we view the judiciary as a rational agent aiming at optimizing a payoff function where economic variables (revenues and costs) and the institutional target of monitoring bribery are linked together. We also characterize the quality of judiciary in terms of her efficiency and integrity (e.g., Caselli and Morelli, 2004).

Technically, we assume the existence of one-to-one correspondences between the oversight activity carried out by J and both the probability $(1 - f)$ with which it

[8] The probabilistic formulation of the objective function is also used by Fredriksson and Svensson (2003) to capture political instability.

finds evidence of the bribe $\widehat{T}$ and the cost, determined by a univariate function S, in terms of effort to be exerted or resource allocation for the anti-corruption task to take place. We thereby consider f as the choice variable for J and denote with $S(f)$ the cost associated with any level of prosecution activity:

$$S(f) = \alpha^{-1}(1-f), \quad \alpha > 0 \tag{3.4}$$

so that identifying any bribe offered to P is costly but bounded from above ($S(0) = 1/\alpha$). The parameter α is a measure of the efficiency of the judiciary; it summarizes the influence of adequate budgetary allocations, sufficient number of staff, adequate training on the judicial work.

The objective function of J reflects the burden of political corruption - in terms of the cost $S(f)$ and the discrepancy (measured by the weighted ex-post transfer $f\widehat{T}$) between the welfare maximizing allocation and the bribery-induced one - as well as the opportunity of benefiting from bribes from the lobby group:

$$V_J(f) = (1-\lambda)C(f) + \lambda\left[-S(f) - f\widehat{T}\right], \quad \lambda \in [0,1]$$

where λ is interpreted as J's level of integrity. $C(f)$ denotes the bribe offered by the lobby group.

For the sake of convenience, the previous expression - taken to be maximized over $f \in [0,1]$ - is written in the following form:

$$V_J(f) = (1+\sigma)^{-1}\left[\sigma C(f) - S(f) - f\widehat{T}\right], \quad \sigma \geq 0 \tag{3.5}$$

where the scalar $\sigma = \lambda^{-1}(1-\lambda)$ can be regarded as the degree of judicial corruptibility.

Under complete and symmetric information, the objectives (3.3) and (3.5) are common knowledge to P, J and the lobby group. To preserve model consistency, we assume that the judiciary needs to find compelling evidence to prosecute corrupt politicians: though bribe payments are observable, bribe-takers can be prosecuted only with probability $(1-f)$[9]. Also, we make a strong assumption in that, if

[9]It is important to emphasize that our conceptual framework deals with illegal rent-seeking. As a consequence, transfers do not take the form of legally enforceable contracts. This evidently creates room for commitment issues since, once the policy is chosen, the lobby group may not feel bound to their promises; however, arguments on political care for future monetary/electoral

identified, the bribe $\widehat{T}$ is confiscated but can neither contribute to financing public goods nor be of any utility to J.

In the following subsection we discuss our modeling choices.

3.3.1 Discussion of the model

Our model can be viewed as a special case of the multiple agents, multiple principles framework developed by Prat and Rustichini (2003) and applied in Fredriksson and Millimet (2007) and Aidt and Hwang (2008)[10]. The assumption that only one group is able to bribe public institutions is important in at least two dimensions. First, it allows for a full characterization of the transfer schedules for general (twice-differentiable) utility functions, as the equilibrium outcome can be implemented by globally truthful transfers. This can fail to be obtained in the case of the multiple lobbies set-up. Second, competition among several lobby groups would typically extend to the different layers of decision-making. Encompassing this case is thus likely to result in considerable complexity, and may require more restrictive assumptions on functional forms for the underlying preferences[11].

We also emphasize that the asymmetric case can be regarded as a good reduced-form representation of political corruption in real-world politics. Jain (1998) argues that corruption differs from legal lobbying in its level of competition: while the latter typically provides transparent rules for potential competitors, bribery is rather a rent-seeking channel through which incumbents prevent potential entrants from entering the bidding process. From this point of view, corruption is commonly described as a more monopsonistic form of rent-seeking (e.g., Lambsdorff, 2002). Furthermore, theoretical work demonstrates that organized groups may decide not to be politically active due to strategic motives (e.g., Aidt, 2002); this idea is supported by some empirical evidence showing that several existing groups are often latent and do not engage in lobbying activities (e.g., Wright 1996).

---

support or career concerns from a repeated-game perspective can be advocated to circumvent this problem.

[10]See also Mazza and van Winden (2008) for an endogenous policy model of a hierarchical government in which multiple policy tasks are shaped by lobbying activities.

[11]On the other side, the presence of multiple lobby groups may instead weaken the policy distortion problem by increasing the costs of rent-seeking and even lead, under some conditions, to the (political) welfare maximizing allocation, the one that would result under absence of lobbying activities (e.g., Mazza and van Winden, 2008).

It is widely recognized that there exist two main dimensions to corruption. On the one side, the supply-side-driven form of corruption is based on the search to extend one's privileges by bribing people in a position of authority and power, for favors that are not consistent with the rule of law. On the other side is demand-side corruption, which is enforced from the top by bribe-takers and does not necessarily rely on the existence of private agents searching for political favors. While both depict important observable facts, they need different modeling assumptions. In the canonical models of corruption (e.g., Shleifer and Vishny, 1993; Aidt and Dutta, 2008), the policy-maker takes the lead and distorts the policy to create rents and attract bribe payments. The common agency framework has been used rather to explain special interest politics in modern democracies, as it seems to fit quite well the political process underlying the formation of economic policies - like public goods provision - where the available policies are bargained upon and exchanged for payments from constituents, that can take the form of campaign support or illegal bribes. The quasi-linear structure of preferences, in which transfers between principal(s) and agent(s) are equivalent to transferable utility, is indeed consistent with the interpretation of lobbying as supply-side-driven corruption (e.g., Coate and Morris, 1999). Moreover, the assumption that the process is driven by bribe-payers helps to model situations in which multiple public decision-makers are subject to influence via sequential bribery, as is the case of our contribution. We recognize, however, that while bribes and influence-buying via campaign contributions and other means cannot be considered perfect substitutes, different assumptions on the way policy-making and lobby groups' activities interact may lead to quite different outcomes and policy implications. This is particularly true in the case of two or more special interest groups. As an example, Epstein and Nitzan (2006) develop a two-tier policy-making model in which rent-seeking government institutions, a legislator and a bureaucrat act as principals of the special interest groups. The authors illustrate how this modeling assumption may distort the tendency of policy-makers to compromise and thus lead to different outcomes with respect to the common agency framework. Nonetheless, in the asymmetric case, the assumption that the bargaining power is fully allocated to the lobby group does not impose any restrictions on the equilibrium outcome of the model. Crucially, the probability f is not taken to be a function of the size of the contribution T, which therefore enters linearly the objectives of the policy-setting authority and of the lobby group; it follows that under Nash bargaining - or any different bargaining structure with full information whose solution is jointly (Pareto) efficient for both the contracting sides - the equilibrium policy proves to be independent of the negotiation process (Grossman and Helpman, 2001).

The specification of the judiciary's objective function deserves further discussion[12]. According to the principles of the rule of law, a leading decision criterion in judicial behavior should be represented by the maximization of the deterrence effect that enforcement measures produce, whereas the judicial system ought to be designed in order to eradicate or circumvent the role of special interests within judges' behavior. A peculiar issue involved in this process is the concern for track records as a measure of career prospects. From a broader perspective, the one that ties the legalist model of judicial behavior and the rational choice theory, the judiciary can be seen as a public decision-maker with a payoff function in which economic incentives, an ethical concern for not fully accomplishing her institutional role, and a negative externality from the expected feedback on career prospects are merged together. Notably, such an approach combines two main aspects, namely rationality - rational agents act to advance their own particular interest - and strategic behavior - all the agents involved in the decision-making process identify the effects of the actual constraints and the interdependence of their actions in a forward-looking way. Thus, the effective set of incentives for judicial decision makers is only partly shaped by their internal structure. In this respect, we stress that our specification does not make the concern for track records conflict with the judicial role of guardian of the rule of law[13].

3.4 Equilibrium analysis

3.4.1 One-layer bribery

We first analyze the case where the lobby group can only bribe the politician,

[12]We thank an anonymous referee for pointing out to us the importance of arguing on this point.

[13]We can thus think of the judge as being penalized in terms of reputation and/or career advancements for having failed to find compelling evidence on the bribes received by the politician.

that is $\sigma = 0$ and (3.5) turns into:

$$V_J(f) = -\left((1-f)\alpha^{-1} + fT(\hat{q})\right), \quad \alpha > 0$$

so that J's objective is given by the sum (with negative sign) of the cost (4) and the weighted ex-post transfer to the politician. The timing of the model is as follows:

(i) J selects the level of oversight activity, determining f, and commits to carry it out;

(ii) the lobby group formulates the bribe schedule $T(q)$;

(iii) P observes $T(q)$ and sets the policy $\hat{q}$; the lobby pays $T(\hat{q})$;

(iv) if not traced, the bribe is received by P.

We derive the subgame-perfect equilibrium (SPE) of the model through backward induction[14]. At the policy-making stage, f^* is predetermined, so using (3.2) and (3.3), the objective function for P becomes:

$$V_P = (f^* - l\theta_1)T(q) + l\sum_{k=1}^{2}\theta_k U_k(q_k)$$

We note that P will give in to the lobby group (accepting $\widehat{T}$) only if the probability of obtaining the bribe exceeds a given threshold (i.e., $l\theta_1$). Indeed, a sufficient condition for the absence of corruption is $l\theta_1 \geq 1$; when the latter is not fulfilled, the lobby group can influence P via bribery only if $f \leq l\theta_1$. More precisely, we have the following:

Lemma 1. *In the SPE of the model:*
i) If $l\theta_1 \geq 1$, no bribery occurs and P chooses $\hat{q}^* = argmax_q l\sum_k \theta_k U_k(q_k)$;
ii) If $l\theta_1 < 1$ and $f^* \leq l\theta_1$, no bribery occurs and P chooses $\hat{q}^*$;
iii) Iff $f^* > l\theta_1$, the lobby group bribes P and obtains $\hat{q}^L > \hat{q}^*$.

In the first two cases, the lobby offers $T = 0$ and no interaction with P takes

[14] The analysis is restricted to the equilibrium profile insofar as the results obtained are invariant with respect to the form of the transfer schedule off-the equilibrium.

place. The equilibrium policy is determined by the first-order condition[15]:

$$\sum_{k=1}^{2} \theta_k U_k'(\hat{q}^*) = \sum_{k=1}^{2} \theta_k n_k G_k'(\hat{q}^*) = 0 \tag{3.6}$$

While resulting in the same equilibrium outcome[16], the first two sub-cases of the Lemma capture two different situations. In case (i), bribing a strongly welfare-oriented politician is not feasible. In case (ii), bribery is feasible but a sufficiently high level of judicial oversight prevents bribery from occurring. We label the first case as First Best (FB) equilibrium, and the second one as Full Deterrence (FD) equilibrium. In the former case, the presence of the judiciary is irrelevant, whereas in the latter, it proves fundamental.

In the third case, the subgame-perfect equilibrium entails corruption. The policy-bribe pair $(\hat{q}^L, T^L)$ is such that $\hat{q}^L$ jointly maximizes the objective function of P and the lobby, the latter acting as a principal. The equilibrium is then defined by

$$(f^* - l\theta_1)\frac{\partial T}{\partial q}(\hat{q}^L) + l\sum_{k=1}^{2} \theta_k U_k'(\hat{q}^L) = 0$$

subject to:

$$U_1'(\hat{q}^L) - \frac{\partial T}{\partial q}(\hat{q}^L) = 0$$

which gives the first-order condition[17]:

$$f^* U_1'(\hat{q}^L) + l\theta_2 U_2'(\hat{q}^L) = 0 \tag{3.7}$$

We note that the main effect of bribery relative to the no corruption case is that the marginal utility of the lobby group gets a larger weight in the political calculus (since $f^* > l\theta_1$). Given the strict concavity of the G_k function, the budget share is biased in favor of the lobby group ($\hat{q}^L > \hat{q}^*$) so that $U_1(\hat{q}^L) > U_1(\hat{q}^*)$ and $U_2(1 - \hat{q}^L) < U_2(1 - \hat{q}^*)$. The equilibrium bribe $T^L \equiv T(\hat{q}^L)$ is such that

$$V_P(\hat{q}^*, 0) = V_P(\hat{q}^L, T^L)$$

[15] This requirement fully characterizes the optimal choice of P, since U_k is concave in $q_1 \equiv q$ for $k = 1, 2$.

[16] In terms of absence of bribery, and hence of the choice of the optimal shares $\hat{q}_k^*$.

[17] Again, sufficiency is guaranteed by concavity.

which delivers:

$$T^L = \frac{l}{f^* - l\theta_1}\left\{\theta_1[U_1(\hat{q}^*) - U_1(\hat{q}^L)] + \theta_2[U_2(\hat{q}^*) - U_2(\hat{q}^L)]\right\} \tag{3.8}$$

Since $\frac{\partial T^L}{\partial q^L} = -\frac{1}{f^* - l\theta_1}\left[l\theta_1 \frac{\partial U_1}{\partial q^L} + l\theta_2 \frac{\partial U_2}{\partial q^L}\right] > 0$ from (3.7), from $\hat{q}^L > \hat{q}^*$ it follows $T^L > 0$. This proves sufficiency for the last claim of Lemma 1. The participation constraint of the lobby group is not binding in equilibrium[18].

Consider now the first stage of the game. In order to perform its anti-corruption task, J selects the level of oversight activity which minimizes $\left[S(f) + f\widehat{T}\right]$ over $f \in [0, 1]$. While $S(f)$ is monotonically decreasing in the control variable, the effect of f on $\widehat{T}$ is ambiguous. Indeed it holds:

$$\frac{\partial \widehat{T}(f)}{\partial f} = -\frac{l}{(f - l\theta_1)^2}\sum_{k=1}^{2}\theta_k\left[U_k(\hat{q}^*) - U_k(\hat{q}^L)\right] - \frac{l}{f - l\theta_1}\sum_{k=1}^{2}\theta_k \frac{\partial U_k}{\partial f} \tag{3.9}$$

This relation captures the equilibrium trade-off induced by a marginal increase in the level of f. The first term on the right-hand side is negative - because a higher f allows the lobby group to lower its bribe schedule for a given f. The second term is positive as it reflects the equilibrium response to bribery. From $\frac{\partial U_1}{\partial f} \equiv \frac{\partial U_1}{\partial q^L}\frac{\partial q^L}{\partial f} > 0$ and $\frac{\partial U_2}{\partial f} \equiv \frac{\partial U_2}{\partial q^L}\frac{\partial q^L}{\partial f} < 0$, and given condition (3.7), the higher the value of f, the larger is the loss of social welfare. As a consequence, the lobby group must offer a larger bribe to compensate P for this loss.

Two different scenarios can therefore emerge. If $l\theta_1 \geq 1$, Lemma 1 ensures that bribery never occurs; with $T = 0$, every non-zero level for the oversight activity of the judiciary is of no consequence, and the optimal choice is $f = 1$[19]. Conversely, if $l\theta_1 < 1$ - a restriction which would be otherwise sufficient for bribery to occur - we can derive a pair of functions $(\hat{q}(f), \widehat{T}(f))$ which map from any value of f in $[0, 1]$ to the corresponding optimal choice of P and the bribe offered by the lobby group, respectively, under which $\widehat{T}(l\theta_1) = 0$ and $\widehat{T}(f) > 0$ for all $f > l\theta_1$.

Given differentiability and compactness assumptions, a solution to J's minimization problem does exist. The following Proposition states that, under a minimum efficiency requirement, the judiciary will be able to fully deter bribery.

[18] That is, $U_1(\hat{q}^L) - T(\hat{q}^L) > U_1(\hat{q}^*)$. This shows that the lobby group has an incentive to bribe P for any $f > l\theta_1$ chosen at the upper node.

[19] This justifies our definition of First Best outcome, since no costly action by the judiciary is necessary to achieve the maximum welfare condition.

Proposition 1. *Let* $l\theta_1 < 1$. *Then there always exists a (finite) threshold* $\overline{\alpha}$ *such that* J *optimally selects* $f^* = l\theta_1$ *if and only if* $\alpha \geq \overline{\alpha}$.

Proof. - See Appendix A.3. □

Hence, in the presence of a sufficiently efficient judiciary, there is no incentive for the lobby group to engage into bribery (Full Deterrence equilibrium). Similarly, full capture (i.e. $f = 1$) results if and only if α is below a given (non-zero) threshold $\underline{\alpha} < \overline{\alpha}$, as any non-zero effort cost for the judicial activity would otherwise reduce her welfare. As a consequence, deterrence effects are hindered by the presence of a highly inefficient (though incorruptible) judiciary: for $\alpha \in (\underline{\alpha}, \overline{\alpha})$, J's optimal choice $f^* \in (l\theta_1, 1)$ results in partial bribery deterrence.

3.4.2 Two-layer bribery

In this section, we investigate the possibility that, while being independent of the political authority, J itself is bribed by the lobby group. With some algebra, it is possible to rewrite the first term on the right-hand side of (3.9) and use (3.7) to obtain:

$$\frac{\partial U_1}{\partial f} - \frac{\partial \widehat{T}}{\partial f} = \frac{1}{(f - l\theta_1)^2}\left\{ l\theta_1[U_1(\hat{q}^*) - U_1(\hat{q}^L)] + l\theta_2[U_2(\hat{q}^*) - U_2(\hat{q}^L)] \right\}$$

which is positive as the payoff of the lobby group is a monotone function of f. This clearly raises the question whether - and under which conditions - it is optimal for the lobby group to bribe the judicial authority at the first stage of the game. When faced with multiple access points to the decision-making, the lobby group needs to evaluate strategically which choices it should attempt to affect, while accounting for the possibility that bribery at one layer may not suffice to fully control reactions at the other.

The objective of the judiciary is now given by equation (3.5), which we repeat below:

$$V_J(f) = (1 + \sigma)^{-1}\left[(-S(f) + \sigma C(f)) - f\widehat{T} \right], \quad \alpha, \sigma > 0 \tag{3.10}$$

whereas P's objective is:

$$V_P = \hat{f}T(q) + l\sum_{k=1}^{2}\theta_k\left\{U_k(q_k) - T_k(q_k) - C_k(\hat{f})\right\} \tag{3.11}$$

where only $T_1 \equiv T$ and $C_1 \equiv C$ can take on non-zero values. The sequence of events therefore includes a preliminary stage where the lobby group decides whether to bribe J via the transfer schedule $C(f)$. J selects the value $\hat{f}$ maximizing (3.10), obtaining the matching monetary reward $C(\hat{f}) \geq 0$. Then, the lobby group decides whether to influence P by submitting the bribe schedule $T(q)$, and finally P chooses a budget allocation $\{\hat{q}, 1-\hat{q}\}$ maximizing (3.11) and collects the bribe $T(\hat{q}) \geq 0$ with probability f.

At the lowest node, P has no power to influence J's choice, and $\hat{f}$ and $C(\hat{f})$ are predetermined. The expression for the optimal $T^L(\hat{f})$ is obtained as in the previous section. At the upper node, the optimal choice of f is jointly efficient for J and the lobby[20]:

$$\hat{f}^L = argmax_{f\in[0,1]}\left\{-S(f) - f\widehat{T}(f) + \sigma C(f)\right\} \quad s.t. \quad \frac{\partial U_1}{\partial f} - \frac{\partial \widehat{T}}{\partial f} - \frac{\partial \widehat{C}}{\partial f} = 0$$

where:

$$C'(f) = \frac{1}{(f-l\theta_1)^2}\left\{l\theta_1[U_1(\hat{q}^*) - U_1(\hat{q}^L)] + l\theta_2[U_2(\hat{q}^*) - U_2(\hat{q}^L)]\right\}$$

is always positive for $f \in (l\theta_1, 1]$. It follows that $\hat{f}^L \geq \hat{f}^*$, with $\hat{f}^*$ denoting the optimal choice for the case with no bribery at J's level.

We can thereby compute $C(\hat{f}^L)$ as the bribe which leaves J indifferent between $\hat{f}^L$ and $\hat{f}^*$:

$$S(\hat{f}^*) + \hat{f}^*\widehat{T}(\hat{f}^*) = S(\hat{f}^L) + \hat{f}^L\widehat{T}(\hat{f}^L) - \sigma C(\hat{f}^L)$$

or:

$$C(\hat{f}^L) = \frac{1}{\sigma}\left[-\frac{(\hat{f}^L - \hat{f}^*)}{\alpha} + \hat{f}^L\widehat{T}(\hat{f}^L) - \hat{f}^*\widehat{T}(\hat{f}^*)\right] \tag{3.12}$$

with the following participation constraint for the lobby group holding in equilib-

[20]Note that bribing the judiciary at this stage of the game is feasible insofar as $\lambda < 1$ ($\sigma > 0$).

rium:

$$U(\hat{q}(\hat{f}^L)) - \widehat{T}(\hat{f}^L) - C(\hat{f}^L) \geq U(\hat{q}(\hat{f}^*)) - \widehat{T}(\hat{f}^*) \tag{3.13}$$

Again, we seek conditions under which, at equilibrium, no bribes to neither P nor J are paid. Hence, we will assume hereafter that $\alpha \geq \overline{\alpha}$[21], so that from Proposition 1 it follows $\hat{f}^* = l\theta_1$. Bribing the judiciary would instead yield $\hat{f}^L$: since $\overline{\alpha}$ has been proven to be the minimal level of efficiency such that $l\theta_1 \equiv argmax\{-S - fT(\hat{q})\}$ in $[0,1]$, and since $C(f) > 0$ for $f > l\theta_1$, we have $\hat{f}^L > l\theta_1$.

The bribe to be paid to J amounts then to:

$$C(\hat{f}^L) = \frac{1}{\sigma}\left[- \frac{(\hat{f}^L - l\theta_1)}{\alpha} + \hat{f}^L \widehat{T}(\hat{f}^L)\right]$$

with $C(\hat{f}^L)$ fulfilling equation (3.13), which is equivalent to requiring:

$$\frac{1}{\sigma}\left[- \frac{(\hat{f}^L - l\theta_1)}{\alpha} + \hat{f}^L \widehat{T}(\hat{f}^L)\right] \leq U(\hat{q}(\hat{f}^L)) - U(\hat{q}^*) - \widehat{T}(\hat{f}^L)$$

Since this holds for any $\alpha \geq \overline{\alpha}$, we state the following:

Proposition 2. *Let $l\theta_1 < 1$. Then there always exists a non-zero threshold $\overline{\sigma}$ such that, for a sufficiently high α, J optimally selects $\hat{f}^* = l\theta_1$ if and only if $\sigma < \overline{\sigma}$.*

Proof. - See Appendix B.3. □

Proposition 2 ensures that, for sufficiently low corruptibility levels, the problem is analogous to that dealt with in section 4.1. A zero-bribe equilibrium - which generalizes our notion of Full Deterrence equilibrium to the case of a corruptible judiciary - is therefore obtained in the two-layer bribery case under identifiable parameter restrictions. In fact, provided that $\sigma < \overline{\sigma}$, a zero-bribe equilibrium is still feasible, conditional on a sufficiently high level of judicial efficiency.

We now characterize this finding in terms of both the efficiency and the integrity of the judiciary (Figure 3.2):

Corollary 1. *Let $l\theta_1 < 1$. Provided that $\sigma < \overline{\sigma}$, there always exists a finite $\overline{\overline{\alpha}}$ such that J optimally selects $f^* = l\theta_1$ if and only if $\alpha \geq \overline{\overline{\alpha}}$. In particular, we have $\overline{\overline{\alpha}} = \alpha(\sigma)$ with $d\overline{\overline{\alpha}}/d\sigma > 0$*

[21]We have already shown that a zero-bribe equilibrium is not feasible if $\alpha < \overline{\alpha}$ even when the judiciary is not corruptible.

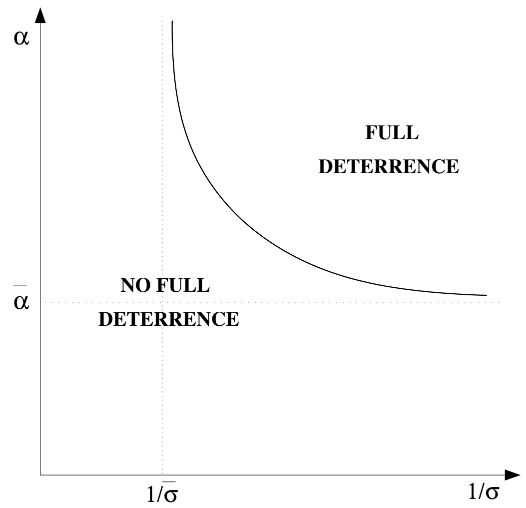

Figure 3.2. Feasibility of the Full Deterrence equilibrium.

Proof. - See Appendix C.3. □

The following claim complements this insight:

Corollary 2. *Let* $l\theta_1 < 1$*. Then* $f = 1$ *obtains if and only if* $\alpha < \underline{\alpha}(\sigma)$ *with* $\underline{\alpha}(\sigma) > 0$ *and* $d\underline{\alpha}/d\sigma > 0$

Proof. - See Appendix D.3. □

3.5 Bribery under judicial dependence

The tension between the independence of judiciaries, the effective oversight and democratic accountability of judges is a well-known concern in the design of institutions and the internal organization of the state. When no other government entity can oversee them, judiciaries enjoy a high level of discretion over choices in their domain and thus may be biased toward those who benefit from a corrupt status quo. Dependent judiciaries, by contrast, are likely to be constrained by politicians who have power over them and thus fail to accomplish their institutional role.

The goal of this section is to explore the relationship between corruption and judicial independence from political influence. We formalize the notion of judicial dependence by assuming that at the beginning of the game the nature (as a pseudo-player) selects all the relevant parameters of the model (i.e. the level of welfare interest of the politician (l), the level of judicial efficiency (α) and the degree of corruptibility of the judiciary (σ)), and yet the politician has the power to change either of the parameters that characterize judicial preferences[22].

We investigate the existence of (possibly partial) deterrence equilbria conditional on judicial dependence. Let us consider the corruptibility of judiciaries, measured by $\sigma > 0$. The case of an independent judiciary, which is immune from political interference, has been developed in the previous section. When l and σ are independently given, the existence space of bribery equilibria is obtained under the requirements of Lemma 1 and Proposition 2. The scenario is depicted in Figure 3, which shows three possible regions according to the thresholds $\overline{\sigma}$ and $\bar{l} = 1/\theta_1$. When $l \geq \bar{l}$ the First Best equilibrium always obtains. If $l < \bar{l}$ but $\sigma < \overline{\sigma}$, Proposition 2 applies and the Full Deterrence equilibrium is achievable conditional on the efficiency of the judiciary being sufficiently high. Only when the welfare interest of P and the integrity of J are both low (that is, $l < \bar{l}$ and $\sigma > \overline{\sigma}$), $\hat{q}^*$ is unfeasible.

Let us now consider the case of a dependent judiciary. In particular we allow the nature to initially choose l and σ, and endow P with the power to select a different J, that is to change the value of σ. We have thereby to consider the possibility that the lobby group bribes P also at this stage of the game to further their political ends. The timing is as follows:

(i) the nature selects l and σ independently;

(ii) the lobby group 1 formulates the bribe schedule $T^I(\sigma)$;

(iii) P chooses either to keep σ or to change it into $\hat{\sigma}$;

(iv) the lobby group 1 formulates the bribe schedule $C(f)$;

(v) J selects $\hat{f}$; the lobby group pays $T^I(\hat{\sigma})$ and $C(\hat{f})$;

(vi) if not traced, $T^I(\hat{\sigma})$ is received by P;

(vii) the lobby group 1 formulates the bribe schedule $T^{II}(q)$;

(viii) P sets the policy $\hat{q}$; the lobby group pays $T^{II}(\hat{q})$;

(ix) if not traced, $T^{II}(\hat{q})$ is received by P.

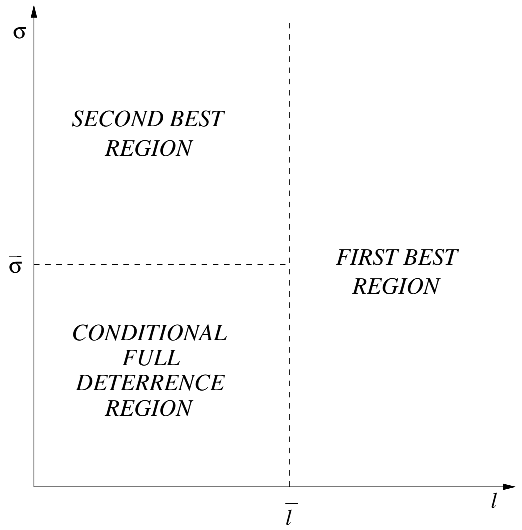

Figure 3.3 Conditional Full Deterrence region.

Note that in V_P it results $T \equiv T^I + T^{II}$. In the last stage of the game the first claim of Lemma 1 still applies and lobbying never occurs provided that $l\theta_1 \geq 1$. As to the solution of the game when $l < 1/\theta_1$, we prove the following:

Proposition 3. *Let $l\theta_1 < 1$. In the SPE of the game:*
i) P chooses the pair $\left\{\hat{\sigma} \to \infty, q^L(\hat{f} = 1)\right\}$;
ii) J selects $\hat{f} = 1$;
iii) lobbying contributions are $\left\{\widehat{C}(f) = 0, \widehat{T}^I(\sigma) = 0, \widehat{T}^{II}(q)\right\}$, where:

$$\widehat{T}^{II}(q) = \frac{l}{1 - l\theta_1} \left\{ \sum_k \theta_k \left[U_k(\hat{q}_k^*) - U_k(\hat{q}_k^L(\hat{f} = 1)) \right] \right\}$$

Proof. - See Appendix E.3. □

Intuitively, the lobby group prefers $f = 1$ to any other $f \in [0, 1)$, since for any $f \leq l\theta_1$, $V_1 = U(\hat{q}^*)$ obtains, while for $f > l\theta_1$ it holds $\frac{\partial V_1}{\partial f} > 0$. For this scenario to emerge, it needs a sufficiently low level of integrity of the judiciary and thus there exists an incentive for the lobby group to influence P at the first stage. Since the

[22]Note that the model remains a game with complete and perfect information in that J and P observe the outcome of the nature's draw.

payoff functions of both the lobby and P are decreasing in C - i.e. the bribe offered to J -, it proves jointly optimal for them to make it null. This turns out to be possible in the case of a corruptible judiciary only if its level of integrity is exactly zero. Hence, in the presence of an incorruptible judiciary, the lobby group and P share a common interest in changing J by letting $\sigma \to \infty$; this result readily follows from observing that $\frac{\partial V_P}{\partial f}$ is positive if $T > 0$ in correspondence of the optimal choice for q.

The equilibrium outcome under judicial dependence is therefore equivalent to that resulting from a society where no such institutional entity exists; in both the cases, in fact, the original finding of Grossman and Helpman (2001) is obtained, according to which the choice of $\hat{q}^L$ is optimally biased in favor of the lobby group. We emphasize that the possibility of achieving the First Best equilibrium is not affected by the structure of the judicial authority. Yet, the presence of the judiciary is redundant only when P is sufficiently welfare-interested and bribery proves unfeasible, since any degree (and form) of institutional dependence will involve the same equilibrium outcome. Importantly, we may observe that, according to Corollary 2, the same conclusion results if P is allowed to determine the level of efficiency of the judiciary, rather than its integrity. In particular, it is easy to show that a (possibly zero) level of efficiency exists for every level of integrity of J, such that whenever P implements it, J optimally selects $f = 1$ and never claims a form of compensation from the lobby group. Both the forms of dependence - organizational as much as hierarchical - appear thereby to be detrimental in the political equilibrium.

It is worth pointing out that the model is consistent with the possibility that the policy-maker could prefer a good judiciary that fights corruption to a corruptible one. In fact, a trade-off exists within the politician's pay-off function, since the presence of corruption creates benefits (bribes) and costs (in terms of political welfare). In the political equilibrium, the two forces balance so that the politician chooses a dependent judiciary only to determine its corruptibility, as its payoff is never inferior when corruption is present. This is a version of the standard result of the literature on endogenous policy, namely that policy-makers do not lose from being captured by special interests. We recognize however that this conclusion depends on the structure of our model and might not extend to other frameworks.

3.6 Conclusion

In this chapter we pursued two different yet related goals: to explore the impact on corruption of the judicial oversight of the political process and, taking advantage of the former, to investigate the relationship between the corruption of public decision-makers (politicians and judges) and the independence of judiciaries. Empirical studies on the topic suggest that the institutional design of judicial authorities is likely to serve as an important determinant of political corruption. Our theoretical framework provides one potential explanation for such evidence. We indeed found that the judiciary plays a critical and heretofore unrecognized role in the shaping of incentives for illegal rent-seeking in policy-making, as it determines the size and the incidence of bribe payments to politicians.

In particular, we showed that, while political corruption can emerge even under fully honest and incorruptible judges, the process of judicial oversight may prevent bribery in the political equilibrium, whenever the judiciary acts in a sufficiently efficient (although corrupt) environment. Efficient judicial systems may in fact countervail lobby groups' influence over prone-to-pressure courts. Hence, provided that a fraction of judges are held not easy to capture, improving the efficiency of judicial systems may serve as a controllable instrument for self-enforced accountability. Only for high levels of corruptibility, a judiciary that is vulnerable to bribery represents an insurmountable impediment to the functioning of the institutional mechanism designed to curb corruption, however well-targeted and efficient, and no deterrence equilibrium is feasible.

We also argued on the role of judicial independence in corrupt societies. Our analysis shows that deterrence equilibria are unfeasible under perfect (hierarchical as well as organizational) dependence on the judiciary of the political power. From a normative perspective, our results suggest that insulating judicial branches from political interference should configure an important issue of institutional design.

Our model delivers several predictions regarding the incidence of corruption in different legal and political systems. First, the location of judges and prosecutors - whether inside the executive branch, autonomous judiciaries or other agencies that are isolated from the regime in power - is a crucial determinant of the degree of judicial oversight of the political process, and thus influences dramatically the occurrence of bribery in the system. Second, states whose institutional arrangements grant the judicial branch strong (functional and hierarchical) independence from political interference should be characterized by weaker correlations between political corruption and judicial corruption, while under judicial dependence the two phenomena should prove highly correlated. However, identifying the causality nexus is not a trivial task: while the presence of strong bribe-providers can generate incentives to make judicial dependence from political power an important aspect of institution design (as predicted by the model), judicial dependence can also prevent the prosecution of political corruption, which thus becomes more attractive and hence more likely (Aaken *et al.*, 2010). Moreover, there seems to be no direct measure for judicial corruption to be exploited in order to derive empirical evidence supporting the model's predictions. We leave this aspect of the analysis to future work.

Appendix to Chapter 3

A.3 Proof of Proposition 1 Denote $G(f) := fT(f)$. First note that from $\widehat{G}(f) = 0$ when $f < l\theta_1$, it follows that J optimizes over $f \in [l\theta_1, 1]$. The Proposition is proven in three steps:

1. Consider α_I such that, for $f \to l\theta_1^+$, $V_J(l\theta_1) > V_J(f)$, that is $S(f) + \widehat{G}(f) > S(l\theta_1)$. Note that the one-sided limits $lim_{f \to l\theta_1} \widehat{G}(f)$ from above and below are finite and equal to zero. The threshold value α_I is accordingly identified through the following second-order Taylor expansion of $\widehat{G}(f)$ around $l\theta_1$:

$$\frac{(1-f)}{\alpha} + \left[\widehat{G}(l\theta_i) + (f_i - l\theta_i)\frac{\partial \widehat{G}}{\partial f} + \frac{(f-l\theta_1)^2}{2}\frac{\partial^2 \widehat{G}}{\partial f^2} + R_{\widehat{G}}\right] > \frac{(1-l\theta_1)}{\alpha}$$

which is equivalent to:

$$(f - l\theta_1)\frac{\partial \widehat{G}}{\partial f}(l\theta_1) + \frac{(f-l\theta_1)^2}{2}\frac{\partial^2 \widehat{G}}{\partial f^2}(l\theta_1) + R_{\widehat{G}} > \frac{(f-l\theta_1)}{\alpha}$$

which in turn holds for:

$$\alpha > \frac{1}{\frac{\partial \widehat{G}}{\partial f}(l\theta_1) + \frac{(f-l\theta_1)^2}{2}\frac{\partial^2 \widehat{G}}{\partial f^2}(l\theta_1) + (f-l\theta_1)^{-1}R_{\widehat{G}}} = \alpha_I$$

2. Let $\hat{f} = argmax\,\{V_J\}$ when $\alpha = \alpha_I$; if $\hat{f} > l\theta_1$, consider α_{II} such that $S(l\theta_1) < S(\hat{f}) + \widehat{G}(\hat{f})$. The previous expression translates into:

$$\frac{(1-l\theta_1)}{\alpha} < \frac{(1-\hat{f})}{\alpha} + \frac{\hat{f}}{\hat{f}-l\theta_1}\left\{l\sum_{k=1}^{2}\theta_k[U_k(q^*) - U_k(q^L)]\right\}$$

which holds for:

$$\alpha > \frac{(\hat{f}-l\theta_1)^2}{\left\{\hat{f}l\sum_{k=1}^{2}\theta_k[U_k(q^*) - U_k(q^L)]\right\}} = \alpha_{II}$$

3. Follow this procedure until at α_N, $\hat{f} = l\theta_1$ obtains. A finite α_N will exist as $\widehat{G}$ is bounded from below (i.e., $\widehat{G} \geq C$ on $(l\theta_1, 1]$, for some constant $C > 0$). We will then have $\overline{\alpha} = \alpha_N$.

The first step ensures V_J has a local maximum at $f = l\theta_1$. The second and third steps ensure this is also the global maximizer in $[l\theta_1, 1]$.

B.3 Proof of Proposition 2 Let $\alpha \to \infty$ and consider the upper bound of the contribution paid by group 1 against $\hat{f}^L$ in this case. We can obtain it by making equation (13) hold with equality (with $\widehat{T}(\hat{f}^*) = 0$ from Proposition 1):

$$\overline{C}(\hat{f}^L) = U(\hat{q}(\hat{f}^L))) - U(\hat{q}(\hat{f}^*)) - \widehat{T}(\hat{f}^L) = \Delta U(\hat{q}(\hat{f}^L)) - \widehat{T}(\hat{f}^L)$$

We can now substitute it in the objective function of J to obtain:

$$\overline{V}_J = -\frac{f}{1+\sigma}\widehat{T}(f) + \frac{\sigma}{1+\sigma}[\Delta U(\hat{q}(f)) - \widehat{T}(f)]$$

where $\overline{V}_J$ represents the maximum utility stream that J could obtain by choosing f. Now we show that there exists a non-zero $\overline{\sigma}$ such that for $\sigma < \overline{\sigma}$ the lobby can never grant J a payoff equal to $\overline{V}_J(l\theta_1)$. Given Lemma 1, we let $f \in [l\theta_1, 1]$ and consider the following:

1. Consider σ_I such that, for $f \to l\theta_1^+$, $\overline{V}_J(l\theta_1) > \overline{V}_J(f)$, that is $-(f+\sigma)\widehat{T}(f) + \sigma\Delta U(\hat{q}(f)) < 0$. We can rewrite this condition by adopting second-order Taylor expansions of $\widehat{G}(f) := f\widehat{T}(f)$, $\widehat{T}(f)$ and $U(\hat{q}(f))$ around $l\theta_1$:

$$-\left[(f-l\theta_1)\left(\frac{\partial\widehat{G}}{\partial f} + \sigma\frac{\partial\widehat{T}}{\partial f}\right) + \frac{(f-l\theta_1)^2}{2}\left(\frac{\partial^2\widehat{G}}{\partial f^2} + \sigma\frac{\partial^2\widehat{T}}{\partial f^2}\right) + R_{\widehat{G}} + \sigma R_{\widehat{T}}\right]$$
$$+\sigma\left[(f-l\theta_1)\frac{\partial U(\hat{q})}{\partial f} + \frac{(f-l\theta_1)^2}{2}\frac{\partial^2 U(\hat{q})}{\partial f^2} + R_U\right] < 0$$

which holds for:

$$\sigma < \frac{(f-l\theta_1)\frac{\partial\widehat{G}}{\partial f} + \frac{(f-l\theta_1)^2}{2}\frac{\partial^2\widehat{G}}{\partial f^2} + R_{\widehat{G}}}{(f-l\theta_1)\left(\frac{\partial U(\hat{q})}{\partial f} - \frac{\partial\widehat{T}}{\partial f}\right) + \frac{(f-l\theta_1)^2}{2}\left(\frac{\partial^2 U(\hat{q})}{\partial f^2} - \frac{\partial\widehat{T}}{\partial f}\right) + \left(R_U - R_{\widehat{T}}\right)} = \sigma_I$$

where all partial derivatives are evaluated at $l\theta_1$.

2. Let $\hat{f} = argmax\,\{V_J\}$ when $\sigma = \sigma_I$; if $\hat{f} > l\theta_1$, consider σ_{II} such that $-(f+$

$\sigma)\widehat{T}(\hat{f}) + \sigma\Delta U(\hat{q}(\hat{f})) < 0$ which holds for:

$$\sigma < \frac{\widehat{G}(\hat{f})}{\Delta U(\hat{q}(\hat{f})) - \widehat{T}(\hat{f})} = \sigma_{II}$$

3. Follow this procedure until at σ_N, $\hat{f} = l\theta_1$ obtains. A finite (non-zero) σ_N will exist since, for $\sigma = 0$, J chooses $\hat{f}^L = l\theta_1$ given $T(l\theta_1) = 0$ and $T(f) > 0$ $\forall f > l\theta_1$, whereas for $\sigma \to \infty$, J chooses $\hat{f}^L > l\theta_1$ as $\Delta U(\hat{q}(f)) - \widehat{T}(f) > 0$ and $\widehat{T}(f)$ is bounded from below. We will then have $\overline{\sigma} = \sigma_N$.

The first step ensures $\overline{V}_J$ has a local maximum at $f = l\theta_1$. The second and third steps ensure this is also the global maximizer in $[l\theta_1, 1]$.

C.3 Proof of Corollary 1 The proof readily follows from Propositions 1 and 2. Indeed, when $\alpha > 0$ is finite, it is easier for the lobby to respect the incentive compatibility constraint of J; in particular we can now rewrite the condition $\overline{V}_J(f) - \overline{V}_J(l\theta_1) = 0$ for $f > l\theta_1$ as:

$$fT(f) - \sigma[\Delta U(\hat{q}(f)) - \widehat{T}(f)] + \Delta S(f) = 0$$

where $\Delta S(f)$ is negative and decreasing in α; so we can obtain the mapping from α to σ of the values that satisfies this expression.

D.3 Proof of Corollary 2 The proof for $\sigma = 0$ is similar to that of Proposition 1. However, in this case we need to show that $f = 1$ maximizes V_J over $[0, 1]$. To reduce notation, let $\frac{\partial G}{\partial f} := G_f$. Define $\alpha_I = 1/G_f(f = 1)$. If $\alpha < \alpha_I$, $G(1)$ is a local maximum for V_J. Now, consider $\hat{f} \equiv argmax\{V_J\} \in [0, 1]$; if $\hat{f} < 1$ define $\alpha_{II} = \frac{1-\hat{f}}{\hat{f}T(\hat{f}) - T(1)} < \alpha_I$. Iterating, we can find $\alpha_N = \underline{\alpha}(0)$ such that $f = 1 \equiv argmax\{V_J\} \in [0, 1]$. Lastly, if for $\alpha = \underline{\alpha}(0)$ $f = 1 \equiv argmax\{V_J\}$, this is true $\forall \alpha < \underline{\alpha}(0)$ since V_J is decreasing in $\alpha \in [0, 1)$.
Now let us consider a generic $\sigma > 0$. We can find $\underline{\alpha}(\sigma)$ with the same procedure as before. Indeed, denoting with $V_J^C = \frac{1}{1+\sigma} V_J^{NC} + \frac{\sigma}{1+\sigma} C$ and V_J^{NC} the payoffs of a corruptible and an incorruptible J respectively, the whole sequence of conditions is easily respected since $C'(f)$ is positive (and the bribe that the lobby group is willing to offer to J is maximum for $f = 1$). Starting from $\alpha_I = \frac{1+\sigma}{(G'(f=1) - \sigma C'(f=1))}$, we can obtain another sequence that converges to $\underline{\alpha}(\sigma)$. Since every term of the sequence

is increasing in σ, it follows that $\underline{\alpha}(\sigma)$ is increasing in σ.

E.3 Proof of Proposition 3 We solve again the game by backward induction:

1. In the third stage, given $\hat{f}$ - that is, the level of control chosen in the second stage -, we obtain $q(\hat{f})$ and $T^{II}(q(\hat{f}))$ as before;

2. In the second stage, given $\hat{\sigma}$, $\hat{f}$ and $C(\hat{f})$ are determined;

3. As to the first stage, we begin determining the optimal solution for the lobby. We previously showed that $\forall f \leq l\theta_1$ $V_1 = U(q^*)$ while, if $f > l\theta_1$, from $\frac{\partial q}{\partial f} > 0$ and $\frac{\partial V_1}{\partial q} > 0$, it follows that $\frac{\partial V_1}{\partial f} > 0$ so that the lobby strictly prefers $f = 1$ to any $f \in [0, 1)$. Also, we can note that both V_1 and V_P are decreasing in C at the optimum. In particular, $V_P^L = l \sum_i \theta_i U_i(q^*) - l\theta_2 C(\hat{f})$ and $V_1^L = U(\hat{q}) - T(\hat{q}) - C(\hat{f})$. From this it follows directly that in the NE of the subgame it is jointly optimal for P and the lobby to set $f = 1$ and $C = 0$. In particular, P chooses $\sigma \to \infty$ since this is equivalent to choosing $f = 1$ with certainty and determining $\widehat{C} = 0$. To show this, we consider the optimal solution for the problem of J. From section 4, $\hat{f}$ maximizes $-(S + fT) + \sigma C$ subject to the constraint $\frac{\partial U}{\partial f} - \frac{\partial T}{\partial f} - \frac{\partial C}{\partial f} = 0$. Then $\hat{f}$ satisfies:

$$-\frac{\partial S}{\partial f} - \hat{f}\frac{\partial T}{\partial f} - T(\hat{f}) + \sigma\left[\frac{\partial U}{\partial f} - \frac{\partial T}{\partial f}\right] = 0$$

or:

$$-\frac{1}{\sigma}\left[\frac{\partial S}{\partial f} - \hat{f}\frac{\partial T}{\partial f} - T(\hat{f})\right] + \frac{\partial U}{\partial f} - \frac{\partial T}{\partial f} = 0$$

so that for $\sigma \to \infty$ the solution to this problem coincides with the optimal choice for the lobby, which was showed to be equal to $f = 1$. Lastly, from (12), we can observe that $C \to 0$ as $\sigma \to \infty$, since the term in brackets is bounded from above.

Accordingly, the whole game reduces to a single stage game where P chooses the pair $(\infty, q^L(\hat{f} = 1))$ and the lobby pays $T^I + T^{II}$, so that in equilibrium it must be:

$$\widehat{T}^I + \widehat{T}^{II} = \frac{1}{1 - l\theta_1}\left\{ l \sum_k \theta_k [U_k(q^*) - U_k(q^L(\hat{f} = 1))] \right\}$$

However, the only time-consistent pair of T^I and T^{II} is given by $T^I = 0$ and $T^{II} = \frac{l}{1-l\theta_1}\Bigg\{\sum_k \theta_k[U_k(q^*) - U_k(q^L(\hat{f} = 1))]\Bigg\}$, since the lobby pays T^I prior to the decision over the policy q^L.

Chapter 4

Lobbying (Strategically Appointed) Bureaucrats

4.1 Introduction

The practice of delegated legislation, according to which bodies other than legislatures are vested with relevant law- and rule-making functions, has a long history in most modern democracies. Delegating policy authority to various bureaucracies involves fundamental issues about policy-making in administrative states. While granting sufficiently high discretion in choice would potentially involve an optimal use of the professionalism and policy expertise of the bureaucracy, it could also encourage independent policy drift and thus require instruments of control (e.g., Gailmard, 2009). The fact that the executive often retains the appointment power can in principle circumvent the agency problem inherent in delegation, yet the conflict between policy goals of higher-level institutions may well exacerbate it (e.g., Epstein and O'Halloran, 1999).

The tension between the value of delegation and the role of interest group influence is also crucial. On the one side, the form of government provides fairly different incentives to influence-seeking activities. On the other, the possibility of lobby-driven bureaucratic drift and the negotiation power of agencies should be taken into account by institutional actors involved in the political process (e.g. Sloof, 2000).

The present study aims to shed some light on the nature of these basic interactions. As a starting point, we exploit Bennedsen and Feldmann (2006)'s model of delegation in which a legislator delegates political authority to a bureaucracy, in order to rely on its policy expertise. An organized group is able to influence the process by initiating bargaining in policy implementation. Rather than viewing the agency subject of influence, i.e. the bureaucrat, as an exogenous entity, we allow it to be strategically appointed by the government administration (i.e. the president or government under separation of powers, or the parliament in a parliamentary system), in light of both the rule of delegation and the policy negotiation process with the organized constituency.

Our main result concerns the critical role of strategic appointments in determining ultimate policy outcomes. It is shown that the impact of bureaucratic lobbying on the allocation of political authority *does not* depend on the preferences and the strength of the interest group in either form of government. In fact, it only involves a strategic response from the authority executing the power of appointment. Hence, the scope of delegation crucially relies on the extent of ideological conflict between the higher-level institutions. As a consequence, the concern for potential policy drift is analogous to that arising in the standard model of legislative delegation without lobbying (e.g. Epstein and O'Halloran, 1999), where it only depends on the relative magnitude of the policy bias with respect to the informational advantage of delegation.

We also provide a comparative investigation of the interest group influence on expected policy outcomes and welfare. We show that lobbying bureaucrats charged with implementing policy plays no role in the political process under both government structures. While bureaucracies are the sole responsible for policy choices under delegated legislation, the actual policy outcomes are (almost) entirely ascribable to the preferences of the higher-level institutions. This finding has important consequences for the theory of agenda setting and political control. First, our analysis contributes to the well-known debate over actual devices available to a legislature to control bureaucratic policy-making, for it suggests that strategic appointments may work as a substitute for legislative oversight and action restrictions (e.g., Gailmard, 2009). Second, from the perspective of optimal statutory design, the model predicts that divided governments should be characterized by more stringent boundaries to which agency decisions ought to conform, independently of the active participation of interest groups in agency decision-making.

To complement our analysis, we present two intuitive modifications of the basic model. The first one reflects the observation that the legal intervention is generally less frequent than appointments of bureaucratic agencies, whereas political systems are typically surrounded by a certain degree of electoral uncertainty. Thus, legislatures may well be unaware of the true preferences of agencies when selecting the amount of policy authority to be delegated. We show that core qualitative findings regarding the impact of bureaucratic lobbying on policy outcomes and their variability remain unaltered. In the second modification, we account for the possibility that the lobby may directly exert pressure to influence the agency selection process, rather than policy implementation[1]. If the government administration is prone to

[1]However, we do not consider situations where the interest group may find it profitable to

pressure from special interests, lobbying *does* have an impact on both the scope of delegation and the expected policy outcome. Compared to the basic model's predictions, this finding suggests that, in the presence of strategic appointments of bureaucracies, mechanisms that hold higher-level institutions accountable for their political role should configure a relevant issue of institutional design.

The remaining chapter is organized as follows. The next section briefly reviews the related literature. Section 3 lays out the basic framework of our analysis, while in section 4 we derive the political equilibria under optimal delegation and strategic agency appointments, and discuss the model's implications regarding the incidence of lobbying across different political systems. Section 5 presents the mentioned modifications of the basic model, while the last section offers concluding remarks. For the sake of exposition, all the proofs and major technical details are relegated to Appendix.

4.2 Related literature

Conceptually, this study is related to several strands of literature. One deals with the optimal allocation of policy authority between elected representatives and non-elected bureaucratic agencies. The focus of a sizable principal-agent literature has been the tradeoff between informational advantages and loss of political control (e.g. Tirole, 1994), and the existence of monitoring abilities and punishment devices for the optimal design of delegation schemes (e.g. Epstein and O'Halloran, 1994, 1999; Huber and Shipan, 2002), while no scope for lobbying is considered. Complementary studies on this issue have investigated instead the monitoring role for interest groups over bureaucrats via an information provision mechanism which activates when legislative policies are intended to serve special interest groups and agencies depart from their legislative mandate (McCubbins and Schwartz, 1984; Banks and Weingast, 1992; Epstein and O'Halloran, 1995).

While numerous recent contributions have paid attention to the subject of policy formation under lobbying in the presence of multilevel (hierarchical) and/or multi-member political structures (e.g. Hoyt and Toma, 1989; Bennedsen and Feldmann,

engage in multi-tier lobbying and thus attempt to influence decision-making at *both* layers (e.g., Mazza and van Winden, 2008).

2002; Epstein and Nitzan, 2002, 2006; Mazza and van Winden, 2008) or the optimality of bureaucratic arrangements - i.e. what should be delegated and what should not - from a social welfare standpoint (Alesina and Tabellini, 2007), relatively few studies have been devoted to the relationship between interest groups influence over decision-making and the potential for delegated legislation. Spiller (1990) develops a multiple principals agency theory to investigate the extent to which legislators could be willing to allocate policy authority to regulators when the latter might be targeted by organized interest groups. Austen-Smith (1993) studies legislative lobbying at the agenda setting stage (Committee) and the voting stage (House), concluding that only agenda stage lobbying is likely to be influential. Diermeier and Myerson (1999) focus on the role of varying constitutional arrangements on the internal organization of legislatures. Their main result concerns the incentives toward the deterrence of collusive behavior that an institutional environment based on the existence of separate (and independent) legislative chambers induces. Sorge (2010) studies the interaction between multi-level lobbying in a divided government and the allocation of political power when multiple policy instruments are available. Similarly, Grajzl (2011) investigates the interplay of political lobbying and delegation exploiting a property rights approach, according to which delegation also involves a rent-dissipation effect from allowing (exclusive) bargaining between the interest group and the bureaucracy.

The present work is distinct from the cited literature in that it accounts for the possibility of strategic agency selection, in the same spirit of Calvert *et al.* (1989), Bertelli and Feldmann (2007) and Krehbiel (2007). Notably, Bertelli and Feldmann (2007) consider a presidential appointment game in which the Senate is required to wield an indirect control over agency decisions by exercising her power of confirmation of presidential nominees and subsequent amendment of implemented policy via direct legislation. We rather incorporate the role of legislatures by acknowledging their power of allocating political authority to various bureaucratic agencies, which in fact are held responsible for a great deal of policy tasks in modern democracies. In doing so, we explicitly address the issue of integrating the theoretical work on interest group lobbying and the rational choice theory of agency selection and delegated legislation from a game-theoretic perspective.

Our analysis is clearly inspired by the seminal work of Bennedsen and Feldmann (2006), who study the effects of bureaucratic lobbying in an otherwise standard imperfect information delegation model. More precisely, their analysis focuses on the influence of interest group lobbying on the bureaucratic policy making and its consequences for optimal statutory design under different political structures. However, their framework does not account for the possibility that the bureaucratic agency be selected strategically, so as to circumvent the effects of bargaining over policy implementation. As a consequence, it does not allow to address the relevant question of whether, and under which conditions, lobbying strategically appointed bureaucrats impacts on the process of legislative delegation and the resulting policy outcome. This work aims at making a step further toward understanding these issues.

4.3 The model

Players and preferences. The model builds upon Bennedsen and Feldmann (2006). Four different players are involved in the political game: an administration (A), a legislator (L), a bureaucrat (B), and an organized interest group (I). The policy space $X \subset \Re$ is one-dimensional, and the policy outcome $x = p + \omega$ is assumed to be a linear function of the effectively chosen policy, p, and of a noise variable, ω, uniformly distributed over $[-r, r]$. We regard r as a measure of the ex ante uncertainty in the political environment, and interpret ω accordingly as specific (unforeseen) contingencies to which new policies are expected to apply.

Factual implementation of policy results from the negotiation between the bureaucratic agency and the special interest group. Under Nash bargaining - or any jointly efficient negotiation rule - the policy outcome always constitutes some compromise between B's and I's ideal policies. Without loss of generality, we restrict attention to the extreme case of take-it-or-leave-it offers from the interest group. As the transfer offered by the lobby enters objectives linearly, this assumption on the allocation of bargaining power does not impose any restrictions on the equilibrium outcome of the model.

All the players have single-peaked preferences over the policy outcome x of the form:

$$U_j(x(p)) = -(x - x_j)^2, \qquad j \in \{A, L\} \tag{4.1}$$

and:

$$U_B(x(p), t) = -(x - x_B)^2 + \alpha_B t(p), \quad \alpha_B > 0 \tag{4.2}$$

$$U_I(x(p), t) = -(x - x_I)^2 - \alpha_I t(p), \quad \alpha_I > 0 \tag{4.3}$$

where $t(p)$ is interpreted as a measure of transferable utility, that the interest group is able to assign to the bureaucracy B. It can be thought of as explicit incentive contract or rather as a promise of future earnings in the private sector (e.g. Grossman and Helpman, 1994, 2001). The parameters (α_I, α_B) are taken to represent the relative values of the transfer from the point of view of the group and the bureaucratic agency, respectively.

Information assumptions. While the actors' ideal points and the objectives (1)-(3) are common knowledge, we maintain the informational rationale of delegation by assuming that the bureaucrat and the interest group only have the expertise to learn the actual realization of the policy shock ω.

Political structure. A distinguishing feature across political systems and forms of government is their connection to political conflict and the extent of executive and legislative power. In the parliamentary system, the constituency of the administration and the legislature are the same, and the fusion of powers is intended to promote the coordination of governmental functions and the implementation of public policies. The principles that inspire the relationships between the branches of government, as derived from the doctrine of the separation of powers, rather view the latter as distinct and independent of one other. We adopt Bennedsen and Feldmann (2006)'s convention to refer to unified government as the parliamentary system and to divided government as the system of separation of powers. Hence, while alignment between administrative and legislative preferences emerges in the former (i.e. $x_A = x_L$), the latter is characterized by a given degree of ideological conflict (i.e. $x_A \neq x_L$). Without loss of generality, we focus on the case $x_L \leq x_A$.

Timing of events. The appointment-delegation-lobbying game unfolds as follows:

Date 0. The administration strategically appoints the bureaucratic agency;

Date 1. The legislator designs a fixed window[2] D, which reflects the scope of delegation, by specifying a reference policy q and a distance $d \geq 0$, such that $D := [q - d, q + d]$;

Date 2. The bureaucrat and the interest group learn the realization of the policy shock ω. Then the interest group formulates an incentive schedule $t(p)$ to be offered

[2]Thus, we only consider the case of a fixed discretion window, as in Bennedsen and Feldmann (2006). The model could be fruitfully generalized to encompassing the possibility of bureaucratic subversion (e.g. Gailmard, 2002).

to the bureaucrat;

Date 3. The bureaucrat selects the policy $p \in D$, and payoffs are realized.

4.4 Optimal delegation and strategic agency appointment

In this section, we derive the subgame perfect equilibrium (SPE) of the model by backward induction. The main results are introduced and discussed separately for the two political systems at issue (in subsections 4.4.1 and 4.4.2 respectively), while conclusions on the degree of policy bias and political influence induced by bureaucratic lobbying will be jointly presented in the last subsection. To ease the comparison of results, we use the superscript P (resp. S) for the parliamentary system (resp. the separation-of-powers system) to label relevant variables and other quantities.

4.4.1 Parliamentary system

The following Proposition characterizes the solution of the appointment-delegation game under bureaucratic lobbying in the parliamentary system (i.e. when $x_A = x_L$).

Proposition 1. *Let* $\beta = \frac{\alpha_B}{\alpha_B + \alpha_I}$ *and* $\hat{x} = \beta x_I + (1 - \beta) x_B^P$. *In the parliamentary system:*

(i) A appoints a bureaucracy whose ideal policy is $x_B^P = (1 - \beta)^{-1}[x_A - \beta x_I]$;

(ii) L chooses the reference policy and the degree of discretion:

$$q^P = x_L, \quad d^P = r \quad \Rightarrow \quad D^P = [x_L - r, x_L + r]$$

(iii) Given (q^P, d^P), *the appointed B is induced by the lobby to implement policy:*

$$p^P = \hat{x} - \omega, \qquad \forall \omega \in [-r, r]$$

under the transfer $t^*(\hat{x}) = \alpha_B (x_I - x_A)^2$.

Proof. - See Appendix A.4. □

Intuitively, the bureaucracy implements the policy yielding the outcome $\hat{x}$ if $|p^P - q^P| \leq d^P$ (delegation window), where $\hat{x}$ has the form of a compromise between B's and I's ideal points in the policy space. The legislature allows for maximum discretion since no bureaucratic drift arises from delegation, given the optimal appointment mechanism at work. Indeed, when appointing the agency, the administration takes into account the induced policy choice p^P that results from the incentive schedule, and thus strategically pins down bureaucratic preferences x_B^P so as to countervail the lobby-driven implementation bias with respect to the no lobbying scenario.

In equilibrium, the legislature fully relies on the bureaucrat's expertise to resolve uncertainty, irrespective of the location of the interest group's ideal point: while the administration is unconstrained in its choice of the agency, the latter enjoys full discretion $d^P = r$ in offsetting any ex post shock ω[3]. In this sense, strategic agency selection works as a perfect substitute for legislative oversight: absent this mechanism, the induced policy outcome would indeed be $\hat{x}' = \beta x_I + (1-\beta)x_B'$, where x_B' denotes the ideal policy of the agent. Hence, only at the knife-edge condition $x_I = x_B'$ - which removes any gain for the interest group from engaging in lobbying - would the legislature grants the bureaucracy full discretion[4].

It is important to emphasize that no lobbying at the implementation stage is not a credible threat, as the interest group will always have an incentive to influence bureaucratic decision-making *even though*, once B is strategically chosen, no deviation of the actual outcome from A's ideal policy x_A can be induced via lobbying[5]. Hence, the equilibrium transfer $t^*(p^P)$ is a strictly positive quantity whenever $x_I \neq x_A$. However, the direction of influence from the interest group is non-influential with respect to the amount of delegated authority, as the mechanism of strategic agency selection prevents the former from moving the bureaucrat's induced policy $\hat{x}$ away from the legislator's ideal one.

The following Corollary summarizes other straightforward results:

Corollary 1. *In the parliamentary system:*

[3]That is, for any $\omega \in [-r, r]$, D^P is symmetrically constructed around L's ideal policy x_L with distance $|x_L - r|$.

[4]Given that the legislature's (expected) welfare is monotonically decreasing in the equilibrium level of $\hat{x}$, it also follows that, without strategic appointments, the legislature experiences a decrease in its utility whenever $|\hat{x}' - x_L| \neq 0$

[5]The lobby's individual rationality constraint is always satisfied for generic (α_I, α_B) (see the Appendix).

(a) *The higher the ex ante uncertainty* r*, the higher the bureaucratic discretion (information rationale);*

(b) *The expected policy outcome with delegation is* $E(x) = x_L$ *(no implementation bias);*

(c) *(i)* $\frac{\partial x_B^P}{\partial r} = 0$*; (ii)* $\frac{\partial x_B^P}{\partial x_A} = \frac{\partial x_B^P}{\partial x_L} > 0$*; (iii)* $\frac{\partial x_B^P}{\partial x_I} < 0$*; (iv)* $\frac{\partial x_B^P}{\partial \beta} > (<)0$ *if* $x_A > (<) x_I$*.*

Proof. - See Appendix B.4. □

Notably, part (*b*) of Corollary 1 underlines that, given the alignment of the administration's preferences with those of the legislature, bureaucratic lobbying does not involve costs for the allocation of policy authority to the bureaucratic agency and no (expected) policy bias - either induced by the lobbying activity or due to independent bureaucratic drift - arises. Hence, legislative delegation allows to rule out any suboptimal outcomes, i.e. those whose distance is such that $|x - x_L| > 0$.

The described equilibrium conditions provide a set of functional relationships upon which we can conduct comparative statics analysis. Not surprisingly, given the absence of ideological conflict between the executive and the legislature, the determination of the $2r$-wide delegation window has no impact on the agency selection, and only the locations of the interest group and administrative's ideal points x_I and x_A, and I's bargaining advantage β[6] determine the optimal choice of the appointee, which moves in the direction that keeps the induced policy outcome $\hat{x}(x_B^P) = x_A$ unaltered.

4.4.2 Separation-of-powers system

A standard result from the theory of delegation without lobbying is that, in a system of separation of powers, the optimal amount of delegated authority is inversely related to the extent of political conflict between the legislature and the executive (Epstein and O'Halloran, 1999).

In the presence of bureaucratic lobbying, Bennedsen and Feldmann (2006) have shown that the impact of interest group influence on the conflict between the legislature's ideal policy and the bureaucratic policy choice, and thus on the scope

[6] We can regard β as bargaining strength in take-it-or-leave-it bargaining in that, *coeteris paribus*, the higher (lower) the relative value α_B (α_I) that the bureaucrat (the interest group) assigns to the transferred resource $t(p)$, the higher β and thus the closer $\hat{x}$ to the lobby's ideal point x_I.

of delegation, is a well-defined function of the location of the lobby. Specifically, Bennedsen and Feldmann (2006) characterize the range of the ideal point x_I for which bureaucratic lobbying brings about an expansion of delegated authority and, *coeteris paribus*, the range for which the legislature delegates more under a system of separation of powers than under a parliamentary structure.

The possibility of strategically selecting the bureaucracy dramatically changes the picture. We show that lobbying at the policy implementation level has no impact on the allocation of decision power to the bureaucracy, which proves only dependent on the wedge between the preferences of the higher-level political institutions and the level of ex ante uncertainty r. This is reflected in the fact that the lobbying-induced policy outcome $\hat{x}$ is a function of neither the lobby's ideal point nor of its bargaining strength, i.e. $\beta x_I + (1-\beta)x_B^S = \hat{x}(x_A, x_L, r)$. The delegation window D^S proves to be a proper (non-degenerate) symmetric interval around x_L, with the policy discretion d^S being a non-linear function of r and the ideological conflict $|x_A - x_L|$. Accordingly, the bureaucracy is not granted maximum discretion in equilibrium.

Importantly, legislative delegation does influence the mechanism of bureaucratic appointment since it shapes the administration's discretion in the agency selection process. The latter proves in fact driven by the policies that appointees can attain once in office. When inducing $\hat{x}$, the administration A needs then to take into account the possibility that the former lies outside the discretion window D^S following legislative delegation. Under this feasibility constraint, the optimal administrative strategy lies in choosing the bureaucrat who implements the constrained policy p^S yielding the outcome $\hat{x}$. This has the main effect of preventing a degenerate delegation window and actually increasing the discretion granted to the bureaucracy (since $x_L < \hat{x} < x_A$).

We now characterize the SPE of the game under separation of powers. First, it is shown how it is generally possible - for any choice of x_B^S and conditional on the realization of ω - that the delegation constraint be binding and corner policies be effectively implemented. Second, a non-zero level of bureaucratic discretion occurs if and only if the (induced) policy bias between the legislature and the bureaucracy does not overtake the size of ex ante policy uncertainty, i.e. iff $|\hat{x} - x_L| < r$. We show that this is always the case. Lastly, the optimal administrative appointee, whose preferences are endogenous to the degree of delegation, is determined strategically.

Proposition 2. *Let x_B^S be the bureaucracy's ideal point as optimally chosen by the* *administrator, and* $\hat{x} = \beta x_I + (1-\beta)x_B^S$ *the induced policy outcome. Also, let* $c := x_A - x_L$ *denote the preference divergence between the administration and the legislature. Then in the separation-of-powers system:*

(i) $\hat{x} > x_L$ *if and only if* $x_A > x_L$*. Furthermore, it holds* $x_L < \hat{x} < x_A$*;*

(ii) A appoints a bureaucracy whose ideal policy is[7]*.:*

$$x_B^S = (1-\beta)^{-1}\left[x_A + \frac{r}{2} - \sqrt{c^2 + \frac{r^2}{4}} - \beta x_I\right]$$

(iii) L chooses the reference policy and the degree of discretion:

$$q^S = x_L, \qquad d^S = r - (\hat{x} - x_L) \quad \Rightarrow \quad D^S = [\hat{x} - r, 2x_L + r - \hat{x}]$$

(iv) Given (q^S, d^S)*, B is induced by the lobby to implement policy:*

$$p^S = \begin{cases} \hat{x} - \omega & \text{if } \omega \in [-r + 2(\hat{x} - x_L), r - 2(\hat{x} - x_L)] \subset [-r, r] \\ 2x_L + r - \hat{x} & \text{if } \omega < -r + 2(\hat{x} - x_L) \\ \hat{x} - r & \text{if } \omega > r - 2(\hat{x} - x_L) \end{cases}$$

under the transfer $t^*(\hat{x}) = \alpha_B^{-1}\left[(\hat{x} - x_B)^2 - (\bar{x} - x_B)^2\right]$*, where:*

$$\bar{x} = \begin{cases} x_B^S & \text{if } \omega \in [-r + 2(\hat{x} - x_L), r - 2(\hat{x} - x_L)] \subset [-r, r] \\ x_L + r - |x_B^S - x_L| + \omega & \text{if } \omega < -r + 2|x_B^S - x_L| \\ x_L - r + |x_B^S - x_L| + \omega & \text{if } \omega > r - 2|x_B^S - x_L| \end{cases}$$

Proof. - See Appendix C.4. □

The main message from Proposition 2 is that, while the bureaucracy's preferences depend (monotonically) on the ideal points of the higher-level political institutions and the value of r, the delegation window is not sensitive to the group's lobbying efforts, insofar as the latter are countervailed by the strategically selected agency. However, from the point of view of the legislature's (expected) welfare, the fact that

[7] If the separation of powers were rather be characterized by the condition $x_L > x_A$, then $\hat{x} < x_L$ and $x_B^S = (1-\beta)^{-1}\left[x_A - \frac{r}{2} + \sqrt{c^2 + \frac{r^2}{4}} - \beta x_I\right]$.

the appointing player exercises political control over bureaucratic choices may prove harmful, if the actual policy outcome is induced more far away from L's ideal point relative to the policy which would prevail absent any strategic selection device[8].

It is relatively straightforward to examine the degree of political influence induced by bureaucratic lobbying across different institutional environments. The following collection of findings will furnish, among other things, a convenient criterion in this regard:

Corollary 2. *In the system of separation of powers:*

(a) $\frac{\partial d^S}{\partial r} > 0$ *(information rationale);*

(b) The expected policy outcome with delegation is $E(x) = x_L + \frac{d}{r}(\hat{x} - x_L)$ *(implementation bias);*

(c) $d^P > d^S$*;*

(d) (i) $\frac{\partial x_B^S}{\partial r} > 0$*; (ii)* $\frac{\partial x_B^S}{\partial x_A} > 0$*; (iii)* $\frac{\partial x_B^S}{\partial x_L} > 0$*; (iv)* $\frac{\partial x_B^S}{\partial x_I} < 0$*; (v)* $\frac{\partial x_B^S}{\partial \beta} > (<) 0$ *if* $\hat{x} > (<) x_I$*.*

Proof. - See Appendix D.4. □

As its analog in Corollary 1, part (a) of the Corollary 2 confirms the role of the informational rationale of legislative delegation, while part (b) demonstrates that allocating decisional power to the bureaucracy necessarily induces bias from the legislature's perspective, independently of the extent of the ex ante uncertainty in the political environment. Nonetheless, benefits from delegation exist, as it prevents outcomes such that $|x - x_L| > |\hat{x} - x_L|$[9].

Contrary to Bennedsen and Feldmann (2006), we thus find that in a system of separation of powers with strategic agency selection, interest group influence neither amplifies nor mitigates the conflict between the legislature's preferences and the policy effectively implemented by the appointed bureaucracy. This finding lies on the ability of the administration to control strategically the agency selection process:

[8] See footnote 3 above and the preceding discussion. Note also that, upon comparing expressions (4) and (7), Proposition 1 reads as a special case of Proposition 2 when $x_A = x_L$.

[9] Given our assumption of uniform distribution for ω, this also means that, from the legislature's standpoint, the distribution of outcomes under optimal delegation (q^S, d^S) first-order stochastically dominates the distribution of outcomes under any other degree of delegation $d^{'} \neq d^S$ (e.g. Bennedsen and Feldmann, 2006).

by appointing the bureaucrat, the executive is able to counteract the impact of the lobby on policy implementation. Hence, the interest group's preferences and bargaining strength do not matter when it comes to the extent of delegation that occurs, regardless of whether the government is unified or divided. As a crucial implication, for given x_L and x_A, the legislature never delegates more to the bureaucracy under a separation of powers system than under a parliamentary structure, independently of the location of the interest group's ideal policy.

Last, basic comparative statics implications are presented for the equilibrium agency selection. A higher level of ex ante uncertainty r, by enlarging the delegation window, allows the administration to appoint an agent further to the right, closer to A's ideal point. The same occurs, *coeteris paribus*, under a shift in the preferences of the administration, whereas point (*iv*) underlines the administrative compensation of a more extreme interest group via the appointment process: if $x_I < x_B^S$, the equilibrium appointee prefers an outcome closer to that preferred by the lobby. Remarkably, the sign of the marginal effect of an increase in the interest group's bargaining strength β depends on the relative location of the induced policy outcome $\hat{x}(x_B^S)$ with respect to the lobby's ideal x_I, and thus also on the ideological conflict between A and L and the size of the ex ante uncertainty r.

Most interesting is perhaps the finding that, although the administrative discretion in agency selection is constrained by legislative delegation, the equilibrium appointee still depends on the administration's most preferred policy outcome x_A. This result relies fully on the presence of information asymmetries between the agents involved in political game: if A were to know in advance the realization $\bar{\omega}$ of the policy shock, then for any $\bar{\omega} \notin [-r + 2c, r - 2c]$ the bureaucrat would be chosen whose ideal policy is such that the induced policy outcome $\hat{x}$ lies at the upper bound of the discretion window, i.e. $\frac{r}{2} - x_L - \bar{\omega}$, which is independent of x_A. In the incomplete information framework, by contrast, the administration makes the bureaucrat's preferences optimally depend on her own ideal policy outcome, at the potential cost of having the appointee implement - upon learning the value of ω - the policy at the boundary of the delegation window.

4.4.3 Alternative government structures and the impact of lobbying

Following Bennedsen and Feldmann (2006), we now turn to contrast the effects of interest group influence in the parliamentary system with those arising under separation of powers. In this regard, two types of measures for the incidence of lobbying are exploited: (1) the expected lobby-induced policy bias under delegation with respect to the case when no lobbying occurs, and (2) the average impact of lobbying on the legislature's expected welfare.

More specifically, we define the lobby's impact on the (expected) policy outcome (LIO):

$$LIO^i = E(x|lobbying) - E(x|no\ lobbying), \quad i \in \{P, S\} \tag{4.4}$$

and the lobby's impact on the (expected) legislature's welfare (LIW):

$$LIW^i = E(U_L(x)|d_l^i, x_l) - E(U_L(x)|d_{nl}^i, x_{nl}), \quad i \in \{P, S\} \tag{4.5}$$

where the subscripts (l, nl) reflect the fact that we are comparing our previous results to the no lobbying scenario. For simplicity, but with no loss of generality, we normalize $x_L = 0$ and thus consider the case $x_A \geq 0$.

In the *parliamentary system*, we have shown that the informational advantage of delegation fully applies and the legislature is willing to rely on the expertise of the bureaucracy by granting the latter the maximum degree of discretion. Under lobbying, $\hat{x} = x_A$ is obtained in equilibrium for any $\omega \in [-r, r]$, and thus $E(x|lobbying) = \hat{x}$ with probability one (see Corollary 1). When the agency is not being lobbied, by contrast, the ally principle $x_A = x_B$ holds[10] and $E(x|no\ lobbying) = x_A$. It readily follows that no lobbying-induced (expected) policy bias arises, i.e. $LIO^P = 0$. Furthermore, given the distribution of outcomes resulting from delegation, the legislature's expected utility from optimal delegation under lobbying simply reads as $E(-(x)^2|d_l^P, \hat{x}) = -\hat{x}^2 = 0$ both with and without lobbying, and hence $LIW^P = 0$.

In the *separation-of-powers system*, delegation always occurs (i.e. $d > 0$) with induced policy bias $\hat{x}(x_B^S) > 0$. The policy outcomes have a simple two-part distribution. Indeed, for $\omega \in [-r+2\hat{x}, r-2\hat{x}]$ the agency selects the policy within its domain, p^S, which yields the outcome $x = \hat{x}$. By contrast, with probability $(1 - d_l^S/r)$, the bureaucrat is constrained by the boundaries of the delegation window and thus - depending on the actual realization of ω - the policy choice varies uniformly in $[-\hat{x}, \hat{x}]$, which extends symmetrically around $x_L = 0$. It follows that the expected policy

[10]Note that, in the presence of interest group influence, the principle of appointing political allies only applies at the knife-edge condition $x_A = x_L = x_I$, as in Bertelli and Feldmann (2007). This powerful result generalizes to our incomplete information framework.

outcome with delegation is simply $E(x|lobbying) = \frac{d_l^S}{r}\hat{x}$. Since lobbying at the implementation level only entails a strategic reaction from the administration - i.e. for $\beta \in (0,1)$, we have $x_{B,l}^S \neq x_{B,nl}^S$ - but not a different induced policy outcome $\hat{x}$ or an alteration in the amount of delegated authority (i.e. $d_l^S = d_{nl}^S$), it is easily shown that $LIO^S = 0$ and $LIW^S = 0$.

We can conclude that the legislature never loses from negotiated bureaucratic policy-making across different political systems, and hence proves indifferent between a biased bureaucracy and an unbiased one, independently of the actual structure of government. On the other hand, the scope of delegation crucially relies on the extent of preference divergence between the legislature and the administration. This result holds true irrespective of how extreme the interest group's preferences are or which degree of bargaining power the latter has. If compared with Bennedsen and Feldmann (2006)'s conclusions on the complex role of bureaucratic lobbying in policy-making, our findings suggest that strategic agency selection is pivotal for the result.

The model also delivers clear implications for the optimal design of statutes in different institutional environments. Notably, political systems or other governance arrangements in which some degree of misalignment between policy goals of the higher-level institutions exists, should be characterized by a lower amount of delegated policy authority. This is in line with the general theory of delegated powers and political control (e.g. Epstein and O'Halloran, 1999).

4.5 Modifications of the basic model

4.5.1 Delegation with unknown bureaucratic preferences

It is well-known that the process of legislation is generally shaped by a wide array of constitutional, statutory and also informal norms, which may severely constrain the timing of law-making (e.g. Gersen and Posner, 2007). As a matter of fact, legal intervention is generally less frequent than appointments of executive agencies and other bureaucratic personnel. Also, democratic political systems are typically surrounded by some degree of electoral uncertainty, and legislature may be uncertain as to which party or coalition will be in government when the delegated policy-forming power is actually exercised. This raises the question of whether and how the costs and benefits of delegated legislation are altered by limited information on the true preferences of delegates. In other words, if the legislature lacks information as to who will receive authority under delegation, which degree of discretion (if any) should the latter be granted?

Here we explicitly address this issue by letting the legislature opt for delegation *before* the administrative design of bureaucratic agencies takes place. More precisely, we consider the following game:

Date 0. The legislator designs the delegation window by specifying the reference policy q and the distance $d \geq 0$;

Date 1. The government administration strategically appoints the bureaucratic agency;

Date 2. The bureaucrat and the interest group learn the realization of the policy shock ω. Then the interest group formulates an incentive schedule $t(p)$ to be offered to the bureaucrat;

Date 3. The bureaucrat selects the policy p, and payoffs are realized.

The solution is easily characterized as follows:

Proposition 3. *Let x_B^j, with $j = \{P, S\}$, be the selected agency in the political system j, and $\hat{x}' = \beta x_I + (1-\beta)x_B^j$. Then:*

(a) In the parliamentary system ($x_A = x_L$):

$$q^P = x_L; \qquad d^P = r; \qquad x_B^P = (1-\beta)^{-1}[x_A - \beta x_I]; \qquad p^P = \hat{x} - \omega$$

(b) In the separation-of-powers system ($x_A - x_L = c > 0$)[11]:

$$q^S = x_L; \qquad d^S = r - |c|; \qquad x_B^S = (1-\beta)^{-1}[x_A - \beta x_I]$$

$$p^S = \begin{cases} x_A - \omega & \text{if } \omega \in [-r+2c, r-2c] \subset [-r, r] \\ x_L + r - c & \text{if } \omega < -r + 2c \\ x_A - r & \text{if } \omega > r - 2c \end{cases}$$

(c) $LIO^j = 0$ and $LIW^j = 0$, $j = \{P, S\}$.

Proof. - See Appendix E.4. □

[11] Again, no loss of generality arises from this simplifying assumption (see footnote 6).

The crucial piece of inference drawn from Proposition 3 is that the core of qualitative findings regarding the impact of bureaucratic lobbying on expected policy outcomes and their variance remains unaltered, independently of the actual form of government. Once again, this confirms the relevance of the appointment process as the primary source of influence on ultimate policy choices. Specifically, as claimed in part (a) and part (b) of the Proposition, the only effect of the reversal of moves is in fact a contraction in the amount of agency discretion - as the legislature attempts to circumvent the (expected) policy bias and the boundaries of the discretion window, upon which the nomination choice is conditioned, become more stringent. This in turn brings about a reduction in the (expected) welfare of the higher-level political institutions. All in all, this result provides a new perspective from which to look at the optimal timing of legal intervention (e.g. Gersen and Posner, 2007), for it suggests a theoretical rationale for timing rules which delay or slow down the legislative process.

4.5.2 Lobbying the administrator

Under strategic appointments, the impact of bureaucratic lobbying on the political process is null. Intuitively, this raises the question of whether the lobby might find it profitable to circumvent her inability to influence policy implementation by interfering directly with the selection of agencies.

In order to analyze the new interactions occurring among the players, we first need to modify our basic framework along two main dimensions. First, while keeping the model's primitives, the administration only is now portrayed as subject of influence. Consistently, A's preferences are expressed by the following single-peaked utility function:

$$U_A(x, t(x)) = -(x - x_A)^2 + \alpha_A t(x), \quad \alpha_A > 0 \tag{4.6}$$

where $t(x)$ and α_A represent the transfer from the lobby - contingent on the bureaucracy's location and thus on the ultimate policy outcome - and the relative value of the latter to A. Assuming no lobbying at B's tier, the bureaucrat's utility function solely depends on its ideal policy x_B, as chosen at the upper stage, that is:

$$U_B(x) = -(x - x_B)^2 \tag{4.7}$$

while the preferences of the legislature and the lobby are characterized as (we set $x_L = 0$ and consider with no loss of generality the case $x_A \geq 0$ and $x_I \neq 0$):

$$U_L(x) = -x^2 \tag{4.8}$$

$$U_L(x, t(x)) = -(x - x_I)^2 - \alpha_I t(x), \quad \alpha_I > 0 \tag{4.9}$$

The second slight adaptation concerns the timing of events. The interest group maintains knowledge of the actual value of the noise parameter ω when lobbying the administrative branch[12]. Once nominated, the bureaucratic agency will in turn learn the realization of the policy shock, owing to its expertise. The new timeline is thus as follows:

Date 0. The interest group learns the realization of ω, then formulates an incentive schedule $t(x)$ to be offered to the administrator;

Date 1. The government administration appoints the bureaucratic agency, that has the expertise to infer ω;

Date 2. The legislator designs the delegation window by specifying the reference policy q and the distance $d \geq 0$;

Date 3. The bureaucrat selects the policy p, and payoffs are realized.

While the analysis of the *administrative lobbying* game is similar to that presented in the previous sections, several differences arise with respect to the basic framework. Crucially, the presence of lobbying at the agency selection stage prevents the administrator from acting strategically. In either political system, therefore, lobbying does affect both the equilibrium level of delegated authority and the (expected) policy outcome. These effects, however, are non-monotonic, as they depend on the relative distance between the lobby's location - and then of the induced policy - and the ideal point of the legislature. In fact, greater agency discretion - both with and without lobbying at A's tier - enlarges the expected policy bias from delegation. Interest group influence can in principle attenuate the conflict between the legislature's preference and the bureaucrat's policy choice - i.e., it might move the expected policy closer to x_L - as a consequence of a reduction in delegation.

The foregoing arguments are summarized in the following set of results:

Proposition 4. *Let* $\lambda := \frac{\alpha_A}{\alpha_A + \alpha_I}$, $\hat{x}'' := \lambda x_I + (1 - \lambda) x_A$, *and* $\Gamma := \frac{r}{2} - \sqrt{x_A^2 + \frac{r^2}{4}}$.

[12] Again, we restrict attention to take-it-or-leave-it offers from the interest group.

Then:

(a) In the political system $j = \{P, S\}$*:*

$$q^j = 0; \qquad d^j = \max\left\{r - |\hat{x}''|, 0\right\}; \qquad x_B^j \equiv \hat{x}''$$

$$p^j = \begin{cases} \hat{x}'' - \omega & \text{if } \omega \in [|\hat{x}''| + \hat{x}'' - r, r - |\hat{x}''| + \hat{x}''] \\ -\frac{r+\omega}{2} & \text{if } \omega < |\hat{x}''| + \hat{x}'' - r \\ \frac{r-\omega}{2} & \text{if } \omega > r - |\hat{x}''| + \hat{x}'' \end{cases}$$

(b) In the parliamentary system ($x_A = 0$)*:*
(i) Administrative lobbying always reduces delegation, i.e. $d_l^P < d_{nl}^P$*;*
(ii) $LIO^P > 0$ *if and only if* $0 < x_I < \frac{r}{\lambda}$*;*
(iii) $LIW^P < 0$*;*

(c) In the separation-of-powers system ($x_A > 0$)*:*
(i) Administrative lobbying increases delegation - i.e. $d_l^S > d_{nl}^S$ *- if and only if:*

$$-\frac{(2-\lambda)x_A}{\lambda} - \frac{\Gamma}{\lambda} < x_I < x_A + \frac{\Gamma}{\lambda}$$

(ii) $LIO^S > 0$ *if and only if:*

$$x_I > -\frac{1-\lambda}{\lambda}x_A, \quad x_I \neq \frac{r}{\lambda} - \frac{1-\lambda}{\lambda}x_A \quad \text{and} \quad \underline{I} < x_I < \overline{I}$$

($\underline{I}$ *and* $\overline{I}$ *given in the Appendix)*
(iii) $LIW^S > 0$ *if and only if* $d_l^S > d_{nl}^A$*;*

(d) The legislature delegates more to the bureaucracy under the separation-of-powers system than under the parliamentary one if and only if:

$$x_I < -\frac{1-\lambda}{2\lambda}x_A$$

Proof. - See Appendix F.4. □

Part (a) of the Proposition is the counterpart to Propositions 2 and 3 in the basic model. It says that the optimal choice of the bureaucrat always emerges as a compromise between the administrator's and the lobby's ideal policies. The optimal level of delegation, which remains a function of the misalignment of preferences of the executive and the legislature, can be higher in either form of government, depending on the relative location of the interest group.

Parts (b) and (c) of the Proposition establish that administrative lobbying always reduces delegation in the parliamentary system - as it biases the appointed bureaucrat away from the legislature's ideal point -, while its effect on optimal delegation in the separation-of-powers system is non-monotonic, depending on the location of the interest group and its bargaining strength. Hence, a higher discretion - that the bureaucrat is granted in exchange for its expertise in reducing political uncertainty - can occur in equilibrium.

Finally, the lobby's preferences and bargaining strength are also crucial to the existence of non-zero effects to both the expected policy bias and the variation in the expected legislature's welfare. Again, equilibrium relationships are nonlinear, given the endogeneity of the appointment mechanism (and lobbying efforts) to the delegation choice. Hence, contrary to the basic model, interest group influence plays here a key role in the political process. This argument makes a strong case for mechanisms that are designed to enforce the accountability of higher-level institutions rather than of bureaucratic agencies.

4.6 Conclusion

The main contribution of this work concerns the analysis of legislative allocation of delegate power and strategic appointments in the presence of bureaucratic lobbying. To this end, it develops a formal model with early-stage strategic agency selection in which the object of choice is the ultimate decision-maker to be vested with political authority, subject to the principles and the rules of delegated legislation. Using this framework, we have showed that bureaucratic lobbying never reduces the scope of delegation across different political systems, as it engenders no influence on the extent of (expected) policy bias induced by delegated legislation. By contrast, the optimal degree of delegated authority emerges as an exclusive (monotonic) relationship between agency discretion, on the one hand, and the ideological conflict between the higher-level institutions and the uncertainty in the political environment, on the other. An important corollary of this result is that the legislature need not shape delegated legislation on the degree of interest group influence on agency decision-making. Rather, the primary source of control on ultimate policy outcomes lies in the process of strategic agency selection. Thus, our analysis raises questions about the wisdom of civil service rules for appointments that preclude government executives from interfering with the nomination of bureaucratic personnel.

Potential applications of the framework presented in this work include the designation of a whole range of bureaucratic entities at various levels, like top bureaucrats, regulatory agencies or other key officials in executive departments, who are entrusted with a (possibly limited) power of rule-making. In this regard, we emphasize that our study is best suited for governance environments where legislation has much to do with the delegation of policy-forming power to bureaucracies. In most political systems, secondary legislation indeed allows bodies other than the legislature to exercise a limited law-making function, when general principles and boundaries are laid down in the primary act. Legislative allocation of delegated authority is in fact a foundational institution in most representatives democracies. From this perspective, the present model may help to better discern the complex processes that generate and govern incentives for the players involved. By contrast, the model's predictions are unlikely to reflect the real working of political systems where the legislative power is almost exclusively vested in the legislature. This may explain the apparent inconsistency that some parliamentary systems which make large recourse to primary law-making, are characterized by lower amounts of delegated policy authority.

Appendix to Chapter 4

A.4 Proof of Proposition 1 Until date 1, the proof is equivalent to that of Lemma 1 in Bennedsen and Feldmann (2006), where the lobby will induce - for any given x_B^P - the bureaucrat to optimally choose the policy outcome $\hat{x} = \beta x_I + (1-\beta)x_B^P$, if the legislative constraint is not binding, i.e. if $(\hat{x} - q) - d \leq \omega \leq (\hat{x} - q) + d$, which is non-empty if and only if $|\hat{x} - q| \leq d + r$. If the bureaucrat is constrained, either of the boundary of the delegation window will be implemented, depending on the realization of ω. Hence, the legislature will optimally set:

$$(q^P, d^P) = argmax_{\{q,d\}}\, EU_L \qquad s.t. \qquad d \geq 0 \tag{4.6}$$

or $q^P = x_L$ and $d^P = r - |\hat{x} - x_L|$ (if the latter is positive, $d^P = 0$ otherwise).

Note that, under knowledge of x_I and β, a one-to-one correspondence exists between the set of choices for the bureaucrat's preferences and the set of policy outcomes. Then, at date 0 the administration will solve:

$$\begin{aligned}
\max_{\hat{x}} \quad EU_A = &-\frac{1}{2r}\int_{-r}^{\hat{x}-x_L-r+|\hat{x}-x_L|}(r - |\hat{x}-x_L| + \omega)^2 d\omega \\
&-\frac{1}{2r}\int_{\hat{x}-x_L-r+|\hat{x}-x_L|}^{\hat{x}-x_L+r-|\hat{x}-x_L|}(\hat{x}-x_L)^2 d\omega \\
&-\frac{1}{2r}\int_{\hat{x}-x_L+r-|\hat{x}-x_L|}^{r}(-r + |\hat{x}-x_L| + \omega)^2 d\omega \\
= &-\frac{1}{r}(\hat{x}-x_L)(r - \frac{1}{2}|\hat{x}-x_L|)
\end{aligned}$$

where we have used the fact that $x_L = x_A$, and $r \geq |\hat{x} - x_L|$. Straightforward maximization leads to $\hat{x} = x_L$ and hence to $x_B^P = (1-\beta)^{-1}[x_A - \beta x_I]$, with $d^P = r$. Accordingly, the legislative constraint is never binding.

It remains to check whether the individual rationality constraint of the interest group is fulfilled, i.e. if the transfer $t^*(\hat{x})$ satisfies:

$$-(\hat{x} - x_I)^2 - \alpha_I t^*(\hat{x}) \geq -(x_B - x_I)^2 \tag{4.7}$$

where the right-hand side term reflect the fact that the interest group is aware of the strategic nomination of B and cannot credibly commit not to lobby the chosen agency. Since the optimal transfer t^* makes the individual rationality constraint of the latter be binding, i.e.:

$$t^*(\hat{x}) = \frac{1}{\alpha_B}(\hat{x} - x_B)^2 = \alpha_B(x_I - x_A)^2$$

the equation (4.7) is verified if and only if $\alpha_B + \alpha_I > 0$, which is always true.

B.4 Proof of Corollary 1 a) Trivial; b) Already shown in subsection 4.4.3; c) Trivial;

C.4 Proof of Proposition 2 Almost the same as Proposition 1, with the exception that at date 0 the administration solves:

$$\begin{aligned}
\max_{\hat{x}} \quad EU_A = &- \frac{1}{2r}\int_{-r}^{\hat{x}-x_L-r+|\hat{x}-x_L|} (r - |\hat{x} - x_L| + \omega - c)^2 d\omega \\
&- \frac{1}{2r}\int_{\hat{x}-x_L-r+|\hat{x}-x_L|}^{\hat{x}-x_L+r-|\hat{x}-x_L|} (\hat{x} - x_L)^2 d\omega \\
&- \frac{1}{2r}\int_{\hat{x}-x_L+r-|\hat{x}-x_L|}^{r} (-r + |\hat{x} - x_L| + \omega - c)^2 d\omega \\
=&\frac{1}{6r}[(-c - |\hat{x}|)^3 - (|\hat{x} - x_L| - c)^3] - \frac{1}{r}(\hat{x} - x_A)^2(r - |\hat{x} - x_L|)
\end{aligned}$$

The first-order condition for this problem is:

$$\begin{aligned}
&\frac{1}{2r}\left[-(-c - |\hat{x}|)^2\frac{|\hat{x} - x_L|}{\hat{x} - x_L} - (|\hat{x} - x_L| - c)^2\frac{|\hat{x} - x_L|}{\hat{x} - x_L}\right] \\
&\quad - \frac{1}{r}\left[(\hat{x} - x_A)(2r - 2|\hat{x} - x_L|) - (\hat{x} - x_A)^2\frac{|\hat{x} - x_L|}{\hat{x} - x_L}\right] = 0
\end{aligned}$$

Assume $\hat{x} > x_L$, then the f.o.c. reduces (after some computation) to:

$$\hat{x}^2 - (2x_A + r)\hat{x} - (x_L^2 - 2x_L x_A - x_A r) = 0 \qquad (4.8)$$

whose real (distinct) solutions are $\hat{x} = x_A + \frac{r}{2} \pm \sqrt{c^2 + \frac{r^2}{4}}$, thus centered around $x_A + \frac{r}{2}$. Hence, the administration will optimal set $x_B^S = (1-\beta)^{-1}[x_A + \frac{r}{2} - \sqrt{c^2 + \frac{r^2}{4}} - \beta x_I]$. Note that $\hat{x} > x_L$ holds if and only if $c > 0$, which is always the case[13]. Also, it

[13]See footnote 7 above for the case in which in the separation-of-powers system we have $c = x_A - x_L < 0$.

clearly obtains that $\hat{x} < x_A$.

The demonstration that the group's IR constraint is not violated, given the equilibrium transfer $t^*(\hat{x})$, is identical to that presented for the parliamentary system as it only relies on the signs of (α_B, α_I).

D.4 Proof of Corollary 2 a) The optimal degree of discretion in the separation-of-powers system is:

$$d^S = \frac{r}{2} - c + \sqrt{c^2 + \frac{r^2}{4}}$$

which is clearly monotonically increasing in r (precisely, $\frac{\partial d^S}{\partial r} \in (0,1)$); b) Already shown in subsection 4.3; c) It follows trivially from $\hat{x} > x_L$; d) Note that we can also write $x_B^S = (1-\beta)^{-1}[r - d(r, x_A, x_L) + x_L - \beta x_I]$. Then:
(i) $\frac{\partial x_B^S}{\partial r} = (1-\beta)^{-1}\left(1 - \frac{\partial d^S}{\partial r}\right) > 0$;
(ii) $\frac{\partial x_B^S}{\partial x_A} = (1-\beta)^{-1}\frac{\sqrt{c^2+\frac{r^2}{4}}-c}{\sqrt{c^2+\frac{r^2}{4}}} > 0$;
(iii) $\frac{\partial x_B^S}{\partial x_L} = (1-\beta)^{-1}\frac{c}{\sqrt{c^2+\frac{r^2}{4}}} > 0$;
(iv) $\frac{\partial x_B^S}{\partial x_I} = -(1-\beta)^{-1}\beta < 0$;
(v) $\frac{\partial x_B^S}{\partial \beta} = (1-\beta)^{-2}(\hat{x} - x_I) > 0$ iff $\hat{x} > x_I$.

E.4 Proof of Proposition 3 For any given (q, d, x_B^j) chosen at the upper node, the agency will implement the policy p^j which leads to $\hat{x}' = \beta x_I + (1-\beta)x_B^j$ if the delegation constraint is not binding, and either $(q+d+\omega)$ or $(q-d+\omega)$ otherwise. Hence, the administration will optimally select x_B^j so as to solve:

$$\begin{aligned}
\max_{\hat{x}'} \quad EU_A = &-\frac{1}{2r}\int_{-r}^{\hat{x}'-q-d}(q+d+\omega-x_A)^2 d\omega \\
&-\frac{1}{2r}\int_{\hat{x}'-q-d}^{\hat{x}'-q+d}(\hat{x}'-x_A)^2 d\omega \\
&-\frac{1}{2r}\int_{\hat{x}'-q+d}^{r}(q-d+\omega-x_A)^2 d\omega \\
=&\frac{1}{r}\left[\frac{1}{3}(r-d)^3 + (q-x_A)^2(r-d) + (\hat{x}'-x_A)^2 d\right]
\end{aligned}$$

which results in $\hat{x}' = x_A$. The legislator will thus design the delegation window by choosing $(q, d \geq 0)$ so as to maximize its expected utility EU_L, i.e. $q^j = x_L$ and $d^j = r - |c|$.Hence, while in the parliamentary system all the previous findings continue to hold, under separation of powers the discretion of the agent is reduced - since $\hat{x} = x_A + \frac{r}{2} - \sqrt{c^2 + \frac{r^2}{4}} < x_A = \hat{x}'$ - and so is the (expected) welfare of the higher-level institutions relative to the case where the bureaucratic preferences are known to the legislature. The last claim of Proposition 3 can be easily proved using the same argument as in subsection 4.4.3 in the main text.

F.4 Proof of Proposition 4 For a given x_B^j, with $j \in \{P, S\}$, and no lobbying at B's tier, the proof is analogous to that of Propositions 1 and 2. The only difference lies in that the delegation variable $d^j \geq 0$ can be zero if $r = |x_B^j|$. Hence, $d^j = \max\left\{r - |x_B^j|, 0\right\}$. The bureaucrat, upon observing ω, chooses $p^j = x_B^j - \omega$ if the delegation constraint is not binding, and either of the boundaries - which are endogenously defined by the reference policy $q^j = x_L = 0$, the level of discretion d^j and the lobby-induced choice of bureaucratic preferences x_B^j - otherwise.

At the first stage, the administrator will appoint the bureaucrat so as to maximize its utility from policy subject to the legislative reaction functions $(q^j(x_B^j), d^j(x_B^j))$ and the incentive $t(x)$ offered by the interest group. Given A's best response, this is equivalent to the following constrained optimization program:

$$\max_{x, t(x)} U_I(x, t(x)) = -(x - x_I)^2 - \alpha_I t(x)$$

s.t.:

$$-(x - x_A)^2 + \alpha_A t(x) \geq -(\tilde{x}^j - x_A)^2, \qquad t(\cdot) \geq 0$$

$$|x - \omega| \leq r - |x|$$

where $\tilde{x}^j$ is the policy outcome the administrator will optimally induce via strategic appointment of B under no lobbying in political system j[14]. The optimal transfer at x is to let the participation constraint be binding, i.e. $t(x) = \alpha A^{-1}[(x - x_A)^2 - (\tilde{x}^j - x_A)^2]$, so that, if the delegation constraint is not binding, the first-order condition yields the unconstrained maximum:

$$\hat{x}'' := \frac{\alpha_A x_I + \alpha_I x_A}{\alpha_A + \alpha_I} = \lambda x_I + (1 - \lambda) x_A = x_B^j, \quad j \in \{P, S\}$$

[14]Hence, $\tilde{x}^P = x_A = 0$ and $\tilde{x}^S = x_A + \frac{r}{2} - \sqrt{x_A + \frac{r^2}{4}}$.

whereas the solution - and hence the bureaucrat's policy choice - corresponds to either of the boundaries of the delegation window if the legislative constraint $D(x_L, r, \hat{x}'')$ holds with equality, which happens if either $\omega < (|\hat{x}''| + \hat{x}'' - r)$ or $\omega > (r - |\hat{x}''| + \hat{x}'')$.

Hence, the optimally chosen bureaucrat is invariant with respect to the actual form of government. However, given the existence of policy conflict in the separation-of-powers system, different implications arise for both the optimal degree of bureaucratic discretion and the impact on expected policies and welfare.

Let us first focus on the parliamentary system. Given $x_A = x_L = 0$, without lobbying the legislature will grant full discretion to the bureaucracy (i.e. $d^P_{nl} = r$) and hence the choice of the bureaucrat will comply with the ally principle (i.e. $x^P_B = x_A$). It readily follows that $d^P_l < d^P_{nl}$.

Also, we have:

$$\begin{aligned} LIO^P = & \frac{d^l}{r}\lambda x_I = \frac{r - \lambda|x_I|}{r}\lambda x_I \\ & >0 \quad iff \quad 0 < x_I < \frac{r}{\lambda} \\ & <0 \quad iff \quad -\frac{r}{\lambda} < x_I < 0 \\ & =0 \quad iff \quad |x_I| = \frac{r}{\lambda} \end{aligned}$$

and $LIW^P < 0$ if and only if $\lambda|x_I| > 0$[15], which is always the case.

In the separation-of-powers system, lobbying increases delegation if and only if:

$$d^S_l > d^S_{nl} \quad \Leftrightarrow \quad |\hat{x}''| < x_A + \Gamma$$

or:

$$-\frac{(2-\lambda)x_A}{\lambda} - \frac{\Gamma}{\lambda} < x_I < x_A + \frac{\Gamma}{\lambda}$$

Note that the upper bound is consistent with the condition $x_I > -\frac{1-\lambda}{\lambda}x_A$, which is necessary and sufficient for $\hat{x}''$ to be positive, and vice versa for the lower bound. Since the legislature's expected welfare is directly related to the equilibrium level of delegation, the same condition on the interest group's location also characterizes the sign of LIW^S, as stated in the main text.

[15]The expected legislature's welfare is indeed a monotonically decreasing decreasing function of $|x^j_B|$, with or without lobbying.

Finally, we have:

$$LIO^S = \frac{d^l}{r}\hat{x}'' - \frac{r - |x_A + \Gamma|}{r}(x_A + \Gamma)$$

and hence:
i) If $x_I \leq -\frac{1-\lambda}{\lambda}x_A$ or $x_I = \pm\frac{r}{\lambda} - \frac{1-\lambda}{\lambda}x_A$, then $LIO^S < 0$;
ii) Let $x_I > -\frac{1-\lambda}{\lambda}x_A$ and $x_I \neq \frac{r}{\lambda} - \frac{1-\lambda}{\lambda}x_A$. Then $LIO^S > 0$ if and only if:

$$[r - \lambda x_I - (1-\lambda)x_A][\lambda x_I + (1-\lambda)x_A] - [r - (x_A + \Gamma)][x_A + \Gamma] > 0$$

which defines the extremes $(\underline{I}, \overline{I})$ of the interval on x_I for which the lobby's impact on expected policy outcome is positive, i.e.:

$$\underline{I} = \Pi(\lambda, x_A, r) - \sqrt{\Pi^2(\lambda, x_A, r) + 4\lambda^2\Omega(\lambda, x_A, r)}$$

$$\overline{I} = \Pi(\lambda, x_A, r) + \sqrt{\Pi^2(\lambda, x_A, r) + 4\lambda^2\Omega(\lambda, x_A, r)}$$

where:

$$\Pi(\lambda, x_A, r) := -\lambda[2(1-\lambda)x_A - r]$$

$$\Omega(\lambda, x_A, r) := \Gamma(\Gamma + 2x_A - r) - \lambda(r - (2-\lambda)x_A)x_A$$

Comparing the optimal level of delegation across the two political systems is straightforward. In fact, we have $d^S > d^P$ if and only if $|\lambda x_I + (1-\lambda)x_A| < \lambda|x_I|$. The latter holds:
i) If $x_I \leq -\frac{1-\lambda}{\lambda}x_A$;
ii) If $x_I > -\frac{1-\lambda}{\lambda}x_A$ and $\lambda x_I + (1-\lambda)x_A < -\lambda x_I$, or $-\frac{1-\lambda}{\lambda}x_A < x_I < -\frac{1-\lambda}{2\lambda}x_A$. This completes the proof.

Part I

References

[1] Abou-Kandil, H., Freilings, G., Jank, G., 1995. On the Solution of Discrete-time Markovian Linear Quadratic Control Problems. *Automatica*, 31, pp. 765-768.

[2] Amman, H.M., Kendrick, D.A., 1999. Linear Quadratic Optimization for Models with Rational Expectations. *Macroeconomic Dynamics*, 3, pp. 534-543.

[3] Amman, H.M., Hendrick, D.A., 2003. Mitigation of the Lucas Critique with Stochastic Control Methods. *Journal of Economic Dynamics and Control*, 27, pp. 2035-2057.

[4] Andersen, T.M., Beier, N.C., 2005. International Transmission of Transitory and Persistent Monetary Shocks Under Imperfect Information, *Journal of International Economics*, 66, pp. 485-507.

[5] Andolfatto, D., Gomme, P., 2003. Monetary Policy Regimes and Beliefs. *International Economic Review*, 44, pp. 1-30.

[6] Aoki, M., Canzoneri, M., 1979. Reduced Forms of Rational Expectations Models. *The Quarterly Journal of Economics*, 93, pp. 59-71.

[7] Azariadis, C., 1981. Self-Fulfilling Prophecies. *Journal of Economic Theory*, 25, pp. 380-396.

[8] Barth, J.R., Swamy, P.A.V.B., Tinsley, P.A., 1982. The Rational Expectations Approach to Economic Modelling. *Journal of Economic Dynamics and Control*, 2, pp. 124-145.

[9] Basar, T., 1989. Some Thoughts on Rational Expectations Models and Alternate Formulations. *Computers and Mathematics with Applications*, 18, pp. 591-604.

[10] Beaudry, P., Portier, F., 2004. An Exploration into Pigou's Theory of Cycles. *Journal of Monetary Economics*, 51, pp. 1183-1216.

[11] Beaudry, P., Portier, F., 2006. Stock Prices, News and Economic Fluctuations. *American Economic Review*, 96, pp. 1293-1307.

[12] Beaudry, P., Portier, F., 2007. When Can Changes in Expectations Cause Business Cycle Fluctuations in Neo-Classical Settings?. *Journal of Economic Theory*, 135, pp. 458-477.

[13] Benhabib, J., Farmer, R.E.A, 1994. Indeterminacy and Increasing Returns. *Journal of Economic Theory*, 63, pp. 19-41.

[14] Benhabib, J., Farmer, R.E.A., 1999. Indeterminacy and Sunspots in Macroeconomics, in: Taylor, J., Woodford, M. (Eds.), Handbook of Macroeconomics, Vol. 1A, North-Holland, Amsterdam, pp. 387-448.

[15] Bertsekas, D.P., 1976. *Dynamic Programming and Stochastic Control*. Academic Press, New York.

[16] Beyer, A., Farmer, R.E.A., 2007. Testing for Indeterminacy: An Application to U.S. Monetary Policy: Comment. *American Economic Review*, 97, pp. 524-529.

[17] Beyer, A., Farmer, R.E.A., 2008. What we Don't Know about the Monetary Transmission Mechanism and why we Don't Know it. *Macroeconomic Dynamics*, 12, pp. 60-74.

[18] Blake, A.P., 1990. The Solution of Time-Varying Linear Rational Expectations Models and the Role of Terminal Conditions. *Economics Letters*, 33, pp. 265-269.

[19] Blair, W.P., Sworder, D.D., 1975. Feedback Control of a Class of Linear Discrete Systems with Jump Parameters and Quadratic Cost Criteria. *International Journal of Control*, 21, pp. 833-841.

[20] Blanchard, O.J., 1979. Backward and Forward Solutions for Economies with Rational Expectations. *American Economic Review*, 69, pp. 114-118.

[21] Blanchard, O.J., Kahn, C., 1980. The Solution of Linear Difference Models under Rational Expectations. *Econometrica*, 48, pp. 1305-1311.

[22] Blanchard, O.J., Fischer, S., 1989. *Lectures On Macroeconomics*. The Mit Press.

[23] Blanchard, O.J., L'Huillier, J.P, Lorenzoni, G., 2009. News, Noise, and Fluctuations: An Empirical Exploration. NBER working paper No. 15015.

[24] Broze, L., Gouriéroux, C., Szafarz, A., 1985. Solutions of Dynamic Linear Rational Expectations Models. *Econometric Theory*, 1, pp. 341-368.

[25] Broze, L., Szafarz, A., 1991. *The Econometric Analysis of Non-Uniqueness in Rational Expectations Models*. Amsterdam.

[26] Bullard, J., Suda, J., 2008. The Stability of Macroeconomic Systems with Bayesian Learners. FRB of St. Louis WP No. 043B.

[27] Carravetta, F., 1996. On the Solution of Linear Stochastic Time-Varying Rational Expectations Models. *Decision and Control, Proceedings of the 35th IEEE*, Volume 1, pp. 55-56.

[28] Carravetta, F., Germani, A., Liptser, R.S., Manes, C., 2002. Filtering of Nonlinear Stochastic Feedback Systems. *SIAM Journal on Control & Optimization*, 40, pp. 1576-1584.

[29] Carravetta, F., Sorge, M.M., 2010. A "Nearly Ideal" Solution to Linear Time-Varying Rational Expectations Models. *Computational Economics*, 35, pp. 331-353.

[30] Cass, D., Shell, K., 1983. Do Sunspots Matter?. *Journal of Political Economy*, 91, pp. 193-227.

[31] Chizeck, H.J., Willsky, A.S., Castanon, D., 1986. Discrete-time Markovian-jump Linear Quadratic Optimal Control. *International Journal of Control*, 43, pp. 213-231.

[32] Chizeck, H.J., Ji, Y., 1988. Optimal Quadratic Control of Jump Linear Systems with Gaussian Noise in Discrete-time, in: *Proceedings of the 27th IEEE Conference on Decision and Control*, pp. 1989-1992.

[33] Chow, G.C., 1980a. Econometric Policy Evaluation and Optimization under Rational Expectations. *Journal of Economic Dynamics and Control*, 2, pp. 47-59.

[34] Chow, G.C., 1980b. Estimation of a Rational Expectations Model. *Journal of Economic Dynamics and Control*, 2, pp. 241-255.

[35] Cochrane, J.H., 1994. Shocks. NBER working paper No. 4698.

[36] Costa, O., Fragoso, M., Marques, R., 2005. *Discrete-Time Markov Jump Linear Systems*, Springer, New York.

[37] Davig, T., 2004. Regime-Switching Debt and Taxation. *Journal of Monetary Economics*, 51, pp. 837-859.

[38] Davig, T., Leeper, E.M., 2007. Generalizing the Taylor Principle. *American Economic Review*, 97, pp. 607-635.

[39] De Santis, A., Germani, A., Scoglio, C., 1993. Kalman Filter Approach to Solution of Rational Expectation Models. *Computers and Mathematics with Applications*, 25, pp. 39-47.

[40] Driskill, R., 2006. Multiple Equilibria in Dynamic Rational Expectations Models: A Critical Review. *European Economic Review*, 50, pp. 171-210.

[41] Dupor, B., Han, J., 2011. Handling Non-Invertibility: Theory and Applications, mimeo.

[42] Evans, G.W., Honkapohja, S., 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton.

[43] Fanelli, L., 2010. Determinacy, Indeterminacy and Dynamic Misspecification in Linear Rational Expectations Models. Quaderni di Dipartimento 4, Department of Statistics, University of Bologna.

[44] Farmer, R.E.A., Waggoner, D.F., Zha, T.A., 2008. Minimal State Variable Solutions to Markov-Switching Rational Expectations Models. *FRB of Atlanta WP No. 2008-23*.

[45] Farmer, R.E.A., Waggoner, D.F., Zha, T.A. 2009. Understanding Markov-switching Rational Expectations Models. *Journal of Economic Theory*, 144, pp. 1849-1867.

[46] Fève, P., Matheron, J., Sahuc, J.-G., 2009. On the Dynamic Implications of News Shocks. *Economics Letters*, 102, pp. 96-98.

[47] Galì, J., 2008. *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press.

[48] Halkin, H., 1974. Necessary Conditions for Optimal Control Problems with Infinite Horizons. *Econometrica*, 42, pp. 267-272.

[49] Hamilton, J.D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, 57, pp. 357-384.

[50] Hamilton, J.D., 1994. *Time Series Analysis*. Princeton University Press, Princeton.

[51] Hansen, L.P., Sargent, T.J., 1980. Formulating and Estimating Dynamic Linear Rational Expectations Models. *Journal of Economic Dynamics and Control*, 2, pp. 7-46.

[52] Jaimovich, N., Rebelo, S., 2009. Can News about the Future Drive the Business Cycle?. *American Economic Review*, 99, pp. 1097-1118.

[53] Kailath, T., 1980. *Linear Systems*. Prentice Hall, Englewood Cliffs.

[54] Kalman, R.E., 1960. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82, pp. 35-45.

[55] Karnizova, L., 2007. News versus Sunspot Shocks in Linear Rational Expectations Models. Department of Economics, University of Ottawa Working Papers, No. 0706E.

[56] Kendrick, D.A., 1981. *Stochastic Control for Economic Models*. McGraw-Hill, New York.

[57] King, R., Watson, M.W., 1998. The Solution of Singular Linear Difference Systems under Rational Expectations. *International Economic Review*, 39, pp. 1015-1026.

[58] King, R., Watson, M.W., 2002. System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations. *Computational Economics*, 20, pp. 57-86.

[59] Klein, P., 2000. Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics and Control*, 24, pp. 1404-1423.

[60] Kriwoluzky, A., 2009. Pre-announcement and Timing - The Effects of a Government Expenditure Shock. EUI Working Paper ECO 2009/40.

[61] Kydland, F.E., Prescott, E.C., 1977. Rules Rather Than Discretion: the Time Inconsistency of Optimal Plans. *Journal of Political Economy*, 85, pp. 473-491.

[62] Leeper, E.M., Walker, T.B., Yang, S.-C.S., 2008. Fiscal Foresight: Analytics and Econometrics. NBER working paper No. 14028.

[63] Leeper, E.M., Walker, T.B., 2011. Information Flows and News Driven Business Cycles. *Review of Economic Dynamics*, 14, pp. 55-71.

[64] Leeper, E.M., Zha, T.A., 2003. Modest Policy Interventions. *Journal of Monetary Economics*, 50, pp. 1673-1700.

[65] Liptser, R.S., Shiryaev, A.N., 2004. *Statistics of Random Processes*. Vol. 2. Berlin, Springer Verlag.

[66] Liu, Z., Waggoner, D.F., Zha, T.A., 2009. Asymmetric Expectation Effects of Regime Shifts and the Great Moderation. *Review of Economic Dynamics*, 12, pp. 284-303.

[67] Lorenzoni, G., 2009. A Theory of Demand Shocks. *American Economic Review*, 99, pp. 2050-2084.

[68] Lubik, T., Schorfheide, F., 2003. Computing Sunspot Equilibria in Linear Rational Expectations Models. *Journal of Economic Dynamics and Control*, 28, pp. 273-285.

[69] Lubik, T., Schorfheide, F., 2004. Testing for Indeterminacy: An Application to U.S. Monetary Policy. *American Economic Review*, 94, pp. 190-217.

[70] Lucas, R.E., 1972. Expectations and the Neutrality of Money. *Journal of Economic Theory*, 4, pp. 103-124.

[71] Lucas, R.E., 1976. Econometric Policy Evaluation. A critique. *Journal of Monetary Economics*, 1, pp. 19-46.

[72] Lucas, R.E., Prescott, E.C., 1971. Investment Under Uncertainty. *Econometrica*, 39, pp. 659-681.

[73] Mankiw, N.G., Reis, R. 2002. Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *The Quarterly Journal of Economics*, 117, pp. 1295-1328.

[74] Marcet, A., Sargent, T.J., 1989. Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models. *Journal of Economic Theory*, 48, pp. 337-368.

[75] McCallum, B.T., 1983. On Non-Uniqueness in Rational Expectations Models: An Attempt at Perspective. *Journal of Monetary Economics*, 11, pp. 139-168.

[76] McCallum, B.T., Nelson, E., 2000.Timeless perspective vs. discretionary monetary policy in forward-looking models. NBER Working Paper No. 7915.

[77] McGough, B., 2003. Statistical Learning with Time-Varying Parameters. *Macroeconomic Dynamics*, 7, pp. 119-139.

[78] Muth, J.F., 1961. Rational Expectations and the Theory of Price Movements, *Econometrica*, 29, pp. 315-335.

[79] Pesaran, M.H., 1989. *The Limits of Rational Expectations*. Oxford: Basil Blackwell.

[80] Rothenberg, T., 1971. Identification in Parametric Models. *Econometrica*, 39, pp. 577-591.

[81] Salemi, M.K., 1986. Solution and Estimation of Linear Rational Expectations Models. *Journal of Econometrics*, 31, pp. 41-66.

[82] Sargent, T.J., 1993. *Bounded Rationality in Macroeconomics*. Oxford: Clarendon Press.

[83] Sargent, T.J., Wallace, N., 1973. The Stability of Models of Money and Growth with Perfect Foresight. *Econometrica*, 41, pp. 1043-1048.

[84] Schmitt-Grohé, S., 1997. Comparing Four Models of Aggregate Fluctuations due to Self-Fulfilling Expectations. *Journal of Economic Theory*, 72, pp. 96-147.

[85] Schmitt-Grohé, S., Uribe, M., 2008. What's News in Business Cycles?. NBER working paper No. 14215.

[86] Shiller, R.J., 1978. Rational Expectations and the Dynamic Structure of Macroeconomic Models: A Critical Review. *Journal of Monetary Economics*, 4, pp. 1-44.

[87] Sims, C., 1980. Macroeconomics and Reality. *Econometrica*, 48, pp. 1-48.

[88] Sims, C., 2002. Solving Linear Rational Expectations Models. *Computational Economics*, 20, pp. 1-20.

[89] Sorge, M.M., 2010a. A Note on Kalman Filter Approach to Solution of Rational Expectations Models. *Economics Bulletin*, 30, pp. 2002-2009.

[90] Sorge, M.M., 2010b. On the Empirical Separability of News Shocks and Sunspots. *Notas Económicas*, 32, pp. 44-55.

[91] Sorge, M.M., 2011a. News Shocks or Parametric Indeterminacy? An Observational Equivalence Result in Linear Rational Expectations Models. *Economics Letters*, in press.

[92] Tucci, M.P., 2004. *The Rational Expectations Hypothesis, Time-Varying Parameters and Adaptive Controls (a Promising Combination?).* Advances in Computational Economics Vol. 19, Springer Verlag.

[93] Wall, K.D., 1980. Generalized Expectations Modeling in Macroeconometrics. *Journal of Economic Dynamics and Control*, 2, pp. 161-184.

[94] Woodford, M., 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

Part II

References

[1] Aaken van, A., Feld, L.P., Voigt, S., 2010. Do Independent Prosecutors Deter Political Corruption? An Empirical Evaluation across Seventy-eight Countries. *American Law and Economics Review*, 12, pp. 204-244.

[2] Ades, A., Di Tella, R., 1997. The New Economics of Corruption: A Survey and some New Results. *Political Studies*, 45, pp. 496-515.

[3] Aidt, T.S., 2002. Strategic Political Participation and Redistribution. *Economics & Politics*, 14, pp. 19-40.

[4] Aidt, T.S., 2003. Economic Analysis of Corruption: A Survey. *Economic Journal*, 113, pp. 632-652.

[5] Aidt, T.S., Dutta, J., 2008. Policy Compromises: Corruption and Regulation in a Democracy. *Economics & Politics*, 20, pp. 335-360.

[6] Aidt, T.S., Hwang, U., 2008. On the Internalization of Cross-National Externalities through Political Markets: The Case of Labour Standards. *Journal of Institutional and Theoretical Economics*, 164, pp. 509-533.

[7] Albanese, G., Sorge, M.M., 2011. The Role of the Judiciary in the Public Decision Making Process. *Economics & Politics*, in press.

[8] Alesina, A., Tabellini, G., 2007. Bureaucrats or Politicians? Part I: A Single Policy Task. *American Economic Review*, 97, pp. 169-179.

[9] Austen-Smith, D., 1993. Information and Influence: Lobbying for Agendas and Votes. *American Journal of Political Science*, 37, pp. 799-833.

[10] Banks, J.S., Weingast, B.R., 1992. The Political Control of Bureaucracies under Asymmetric Information. *American Journal of Political Science*, 36, pp. 509-516.

[11] Becker, G.S., Stigler, G.J., 1974. Law enforcement, Malfeasance, and the Compensation of Enforcers. *Journal of Legal Studies*, 3, pp. 1-18.

[12] Bennedsen, M., Feldmann, S.E., 2006. Lobbying Bureaucrats. *Scandinavian Journal of Economics*, 108, pp. 643-668.

[13] Bertelli, A.M., Feldmann, S.E., 2007. Strategic Appointments. *Journal of Public Administration Research and Theory*, 17, pp. 19-38.

[14] Bond, P., 2009. Contracting in the Presence of Judicial Agency. *The B.E. Journal of Theoretical Economics*, 9, Article 36.

[15] Boudreaux, D.J., Pritchard, A.C., 1994. Reassessing the Role of the Independent Judiciary in Enforcing Interest-Group Bargains. *Constitutional Political Economy*, 5, pp. 1-21.

[16] Buchanan, J.M., Tullock, G., 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. University of Michigan Press.

[17] Calvert, R.L., McCubbins, M.D., Weingast, B.R., 1989. A Theory of Political Control and Agency Discretion. *American Journal of Political Science*, 33, pp. 588-611.

[18] Caselli, F., Morelli, M., 2004. Bad Politicians. *Journal of Public Economics*, 88, pp. 759-782.

[19] Coate, S., Morris, S., 1999. Policy Persistence. *American Economic Review*, 89, pp. 1327-1336.

[20] Dakolias, M., 1999. Court Performance Around the World: A Comparative Perspective. *Yale Human Rights and Development Law Journal*, 2, pp. 87-142.

[21] Dal Bó, E., Dal Bó, P., Di Tella, R., 2006. *'Plata o Plomo?'*: Bribe and Punishment in a Theory of Political Influence. *American Political Science Review*, 100, pp. 1-13.

[22] Diermeier, D., Myerson,R.B., 1999. Bicameralism and Its Consequences for the Internal Organization of Legislatures. *American Economic Review*, 89, pp. 1182-96.

[23] Dixit, A., Grossman, G.M., Helpman, E., 1997. Common Agency and Coordination: General Theory and Application to Government Policy Making. *Journal of Political Economy*, 105, pp. 752-769.

[24] Djankov, S., La Porta, R., Lopez-de-Silane, F., Shleifer, A., 2002. Courts: the Lex Mundi Project. NBER Working Papers No. 8890.

[25] Epstein, D., O'Halloran, S., 1994. Administrative Procedures, Information, and Agency Discretion. *American Journal of Political Science* 38, pp. 697-722.

[26] Epstein, D., O'Halloran, S., 1995. A Theory of Strategic Oversight: Congress, Lobbyists, and the Bureaucracy. *Journal of Law, Economics, and Organization*, 11, pp. 227-255.

[27] Epstein, D., O'Halloran, S., 1999. *Delegating Powers: A Transaction Cost Politics Approach to Policy Making Under Separate Powers*. Cambridge: Cambridge University Press.

[28] Epstein, G.S., Nitzan, S., 2002. Endogenous Public Policy, Politicization and Welfare. *Journal of Public Economic Theory*, 4, pp. 661-677.

[29] Epstein, G.S., Nitzan, S., 2006. Lobbying and Compromise. *Public Choice*, 126, pp. 317-330.

[30] Feld, L.P., Voigt, S., 2003. Economic Growth and Judicial Independence: Cross Country Evidence Using a New Set of Indicators. *European Journal of Political Economy*, 19, pp. 497-527.

[31] Feld, L.P., Voigt, S.. 2004. Making Judges Independent. Some Proposals Regarding the Judiciary. CESifo working paper No. 1260.

[32] Fredriksson, P., Millimet, D., 2007. Legislative Organization and Pollution Taxation. *Public Choice*, 131, pp. 217-242.

[33] Fredriksson, P., Svensson, J., 2003. Political instability, corruption and policy formation: the case of environmental policy. *Journal of Public Economics*, 87, pp. 1383-1405.

[34] Gailmard, S., 2002. Expertise, Subversion, and Bureaucratic Discretion. *Journal of Law, Economics, and Organization*, 18, pp. 536-555.

[35] Gailmard, S., 2009. Discretion Rather than Rules: Choice of Instruments to Control Bureaucratic Policy Making. *Political Analysis*, 17, pp. 25-44.

[36] Gersen, J.E., Posner, E.A., 2007. Timing Rules and Legal Institutions. *Harvard Law Review*, 121, pp. 543.

[37] Glaeser, E.L., Shleifer, A., 2002. Legal Origins. *The Quarterly Journal of Economics*, 117, pp. 1193-1229.

[38] Grajzl, P., 2011. A Property Rights Approach to Legislative Delegation. *Economics of Governance*, 12, pp. 177-200.

[39] Grossman, G.M., Helpman, E., 1994. Protection for Sale. *American Economic Review*, 84, pp. 833-850.

[40] Grossman, G.M., Helpman, E., 2001. *Special Interest Politics.* Cambridge MA., MIT Press.

[41] Hanssen, A.F., 2004. Is There a Politically Optimal Level of Judicial Independence?. *American Economic Review*, 94, pp. 712-729.

[42] Harstad, B., Svensson, J., 2011. Bribes, Lobbying and Development. *American Political Science Review*, 105, pp. 46-63.

[43] Hellman, J.S., Jones, G., Kaufman, D., 2000. Seize the State, Seize the Day: State Capture, Corruption, and Influence in Transition. World Bank Policy Research Paper No. 2444.

[44] Hillman, A.L., 1982. Declining Industries and Political-Support Protectionist Motives. *American Economic Review*, 72, pp. 1180-1187.

[45] Hillman, A.L., 1989. *The Political Economy of Protection.* Chur: Harwood.

[46] Hillman, A.L., Katz, E., 1987. Hierarchical Structure and the Social Cost of Bribes and Transfers. *Journal of Public Economics*, 34, pp. 129-142.

[47] Hoeflich, M.H., 1984. Regulation of Judicial Misconduct from Late Antiquity to the Early Middle Ages. *Law and History Review*, 2, pp. 79-104.

[48] Hoyt, W.H., Toma, E.F., 1989. State Mandates and Interest Group Lobbying. *Journal of Public Economics*, 38, pp. 199-213.

[49] Huber, J.D., Shipan, C.R., 2002. *Deliberate Discretion? The Institutional Foundations of Bureaucratic Autonomy*. Cambridge: Cambridge University Press.

[50] Immordino, G., Pagano, M., 2010. Legal Standards, Enforcement and Corruption. *Journal of the European Economic Association*, 8, pp. 1104-1132.

[51] Krehbiel, K., 2007. Supreme Court Appointments as a Move-the-Median Game. *American Journal of Political Science*, 51, pp. 231-240.

[52] Jain, A.K., 1998. Models of corruption, in A. K. Jain (ed.), *Economics of Corruption. Recent Economic Thought Series*, 65, Boston and London: Kluwer Academic.

[53] Jain, A.K., 2001. Corruption: A review. *Journal of Economic Surveys*, 15, pp. 71-121.

[54] La Porta, R., Lopez-de-Silanes, F., Pop-Eleches, C., Shleifer, A., 2004. Judicial Checks and Balances. *Journal of Political Economy*, 112, pp. 445-470.

[55] Lambsdorff, J.G., 2002. Making Corrupt Deals: Contracting in the Shadow of Law. *Journal of Economic Behavior and Organization*, 48, pp. 221-241.

[56] Landes, W.M., Posner, R.A., 1975. The Independent Judiciary in an Interest-Group Perspective. *Journal of Law&Economics*, 18, pp. 875-901.

[57] Mauro, P., 1995. Corruption and Growth. *The Quarterly Journal of Economics*, 110, pp. 681-712.

[58] Mazza, I., van Winden, F., 2008. An Endogenous Policy Model of Hierarchical Government. *European Economic Review*, 52, pp. 133-149.

[59] McCubbins, M.D., Schwartz, T., 1984. Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science*, 28, pp. 165-179.

[60] Olson, M., 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge: Harvard University Press.

[61] Polinsky, A.M., Shavell, S., 2001. Corruption and Optimal Law Enforcement. *Journal of Public Economics*, 81, pp. 1-24.

[62] Posner, R.A., 1994. What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does). *Supreme Court Economic Review*, 3, pp. 1-41.

[63] Posner, R.A., 1995. *Overcoming Law*. Cambridge MA.: Harvard University Press.

[64] Priks, M., 2011. Judiciaries in Corrupt Societies. *Economics of Governance* 12, 75-88.

[65] Rose-Ackerman, S., 1978. *Corruption: A Study in Political Economy*. New York: Academic Press.

[66] Rubin, P.H., Curran, C., Curran, J.F., 2001. Litigation Versus Legislation: Forum Shopping by Rent Seekers. *Public Choice*, 107, pp. 295-310.

[67] Shleifer, A., Vishny, R.W., 1993. Corruption. *The Quarterly Journal of Economics*, 108, pp. 599-617.

[68] Sloof, R., 2000. Interest Group Lobbying and the Delegation of Policy Authority. *Economics & Politics*,12, pp. 247-274.

[69] Sorge, M.M., 2010c. Lobbying-consistent Delegation and Sequential Policy Making. *Economics Bulletin*, 30, pp. 3088-3102.

[70] Sorge, M.M., 2011b. Lobbying (Strategically Appointed) Bureaucrats. Mimeo

[71] Spiller, P.T., 1990. Politicians, Interest Groups, and Regulators: A Multiple-Principals Agency Theory of Regulation, or 'Let Them be Bribed'. *Journal of Law and Economics*, 33, pp. 65-101.

[72] Stigler, G.J., 1971. The Theory of Economic Regulation. *The Bell Journal of Economics*, 2, pp. 359-365.

[73] Tirole, J., 1994. The Internal Organization of Government. *Oxford Economic Papers*, 46, pp. 1-29.

[74] Wright, J.R., 1996. *Interest Groups and Congress*, Boston: Allyn & Bacon.