

# SET IDENTIFIED DYNAMIC ECONOMIES AND ROBUSTNESS TO MISSPECIFICATION

ANDREAS TRYPHONIDES  
HUMBOLDT UNIVERSITY

**ABSTRACT.** We propose a new inferential methodology for dynamic economies that is robust to misspecification of the mechanism generating frictions. Economies with frictions are treated as perturbations of a frictionless economy that are consistent with a variety of mechanisms. We derive a representation for the law of motion for such economies and we characterize parameter set identification. We derive a link from model aggregate predictions to distributional information contained in qualitative survey data and specify conditions under which the identified set is refined. The latter is used to semi-parametrically estimate distortions due to frictions in macroeconomic variables. Based on these estimates, we propose a novel test for complete models. Using survey data collected by the European Commission, we apply our method to estimate distortions due to financial frictions in the Spanish economy. We investigate the implications of these estimates for the adequacy of the standard model of financial frictions SW-BGG ([Smets and Wouters \(2007\)](#), [Bernanke, Gertler, and Gilchrist \(1999\)](#)).

Keywords: Frictions, Heterogeneity, Set Identification, Surveys, Testing

JEL Classification: C13, C51, C52, E10, E60

---

*Date:* 5/01/2018.

This paper is based on chapter 2 of my PhD thesis (European University Institute). I thank Fabio Canova for his advice and help with earlier versions of this paper, and the rest of the thesis committee: Peter Reinhard Hansen, Frank Schorfheide and Giuseppe Ragusa for comments and suggestions. This paper benefited also from discussions with George Marios Angeletos, Eric Leeper, Ellen McGrattan, Juan J. Dolado, Thierry Magnac, Manuel Arellano, Eleni Aristodemou, Raffaella Giacomini, and comments from participants at the 2nd IAAE conference, the 30th EEA Congress, T2M -BdF workshop, the ES Winter Meetings, the seminars at the University of Cyprus, Royal Holloway, Humboldt University, Universidad Carlos III, University of Glasgow, University of Southampton, the Econometrics and Applied Macro working groups at the EUI. Any errors are my own.

## 1. INTRODUCTION

Incorporating frictions in dynamic macroeconomic models is undoubtedly essential, both in terms of their theoretical and empirical implications. They are being routinely used to explain both the observed persistence of macroeconomic time series and several stylized facts obtained from firm and household level data. However, at the macroeconomic level, the selection between different types of frictions and the specification of their corresponding mechanism is a complicated process featuring arbitrary aspects. This arbitrariness implies that different studies may find support for alternative types of frictions, depending on other assumptions made. For example, the choice of nominal rigidities might depend on which real rigidities are included in the model or whether there is variable capital utilization or not<sup>1</sup>. More importantly, unless frictions are micro-founded, policy conclusions may become whimsical as different mechanisms imply different relationships between policy parameters and economic outcomes.

In this paper we propose a new inferential methodology that is robust to misspecification of the mechanism generating frictions in a dynamic stochastic economy. The approach treats economies with frictions as perturbations of a frictionless economy, which are not uniquely pinned down, and are consistent with different structural specifications.

The paper makes several contributions. First, we derive a characterization for the law of motion of an economy with frictions that imposes identifying restrictions on the solution of the model. [Chari, Kehoe, and McGrattan \(2007\)](#), CKM hereafter, identify wedges in the optimality conditions of a frictionless model that produce the same equilibrium allocations in economies with specific parametric choices for the

---

<sup>1</sup>See the analysis of [Christiano, Eichenbaum, and Evans \(2005\)](#) and their comparison to the results of [Chari, Kehoe, and McGrattan \(2000\)](#).

frictions. We also take a frictionless model as a benchmark but contrary to CKM we construct a general representation for the wedge in the decision rules. We illustrate through examples that the sign of the conditional mean of the decision rule wedge is typically known, even when the exact mechanism generating the friction is unknown. We utilize knowledge of the sign of the conditional mean to set-identify the parameters of the benchmark model.

Moment inequality restrictions have been used to characterize frictions in specific markets, see for example [Luttmer \(1996\)](#) and [Chetty \(2012\)](#). To our knowledge, we are the first to characterize such restrictions in dynamic stochastic macroeconomic models and to show their relationship with the literature on wedges in equilibrium models. We are also the first to characterize set identification in this class of models. The econometric theory we develop can accommodate inequality restrictions of general form.

Due to set identification, many models with frictions are likely to be consistent with the robust identifying restrictions. Thus, additional data other than macroeconomic time series can be potentially useful in order to further constrain the set of admissible models. As a second contribution, our methodology permits the introduction of additional restrictions. We show how qualitative survey data can be useful for this purpose, since they provide distributional information which is highly relevant in models featuring frictions and heterogeneity. More specifically, qualitative survey data are linked to current and future beliefs of agents in the model. If survey data reflect subjective conditional expectations, they contain important indications about agents' actions through the behavioral equations. In the literature<sup>2</sup>, survey data are typically linked to model based random variables using additional observation equations with

---

<sup>2</sup>We mainly refer to the treatment of survey data in the most recent "modern" DSGE literature. The treatment of survey data in Rational Expectations models date back to the work of [Pesaran \(1987\)](#) and others.

additive measurement error (see e.g. [Del Negro and Schorfheide \(2013\)](#)). However, this paper follows a different approach as the model is incomplete and there is no well defined model quantity that can be linked to the data. This is where the law of motion representation we derive proves useful, as we can link the macroeconomic distortions to aggregated qualitative surveys through additional moment inequality restrictions.

Regarding possible applications of the methodology, we show how it contributes to model selection, where the object of interest is not the set-identified parameter vector itself, but the implied semi-parametric estimates of frictions. To that end we propose a novel Wald test that compares the estimates of distortions implied by a candidate model to the robust set obtained using our method. We also derive large sample theory for this type of test. Beyond restricting the set of observationally equivalent distortions, the use of additional data i.e. surveys, prevents the test statistic from degenerating when the data generating process is such that the incomplete model nests the candidate model.

We apply the methodology to estimate the distortions present in the Spanish economy due to financial frictions. We estimate a small open economy version of the [Smets and Wouters \(2007\)](#) model using qualitative survey data collected by the Commission on the financial constraints of firms. We contrast the implied estimates of distortions to macroeconomic aggregates due to these frictions to those identified using a complete model that incorporates the [Bernanke, Gertler, and Gilchrist \(1999\)](#) financial accelerator.

We have already referred to how our paper relates to some strands of literature, namely the literature on wedges and frictions (i.e. [Chari, Kehoe, and McGrattan \(2007\)](#), [Christiano, Eichenbaum, and Evans \(2005\)](#), [Chari, Kehoe, and McGrattan \(2000\)](#)) and the literature on including survey data in DSGE models (i.e. [Del Negro](#)

and Schorfheide (2013)). We contribute to the literature that deals with partial identification in structural macroeconomic models (e.g. Lubik and Schorfheide (2004); Coroneo, Corradi, and Santos Monteiro (2011)) and the literature on applications of moment inequality models, see Pakes, Porter, Ho, and Ishii (2015) and references therein. We postpone illustrating more detailed connections to the vast partial identification literature for later sections, when we deal with estimation and inference using our methodology.

The rest of the paper is organized as follows. Section 2 provides a motivating example of the methodology in a partial equilibrium context, while section 3 presents the prototype economy which will be used as an experimental lab to illustrate our methodology. Section 4 examines the distortions present in the decision rules and their observable aggregate implications. Section 5 provides the formal treatment of identification without additional information while Section 6 analyzes the case with additional information, including qualitative survey data. Section 7 discusses inference issues and provides the test. Section 8 contains the application to Spanish data and Section 9 concludes and provides avenues for future work. Appendix A contains proofs and the empirical results. Appendix B (online) contains more examples of moment inequality restrictions, details on the relation to the model uncertainty literature and the algorithm to compute frictions, an example on identification using our method, an illustration of the validity of bootstrap for the proposed test and a plot of the survey data used in the application.

Throughout the paper we refer to three different conditional probability measures. The objective probability measure,  $\mathbb{P}_t$ , the probability measure determined by the frictionless model  $M_f$ ,  $\mathbb{P}_t^{M_f}(\cdot|\cdot)$ , and the subjective probability measure  $\mathcal{P}_{t,i}$  where  $i$  identifies agent  $i$  and  $t$  the timing of the conditioning set. All three are absolutely

continuous with respect to Lebesgue measure. The corresponding conditional expectation operators are  $\mathbb{E}_t(\cdot)$ ,  $\mathbb{E}_t^{M_f}(\cdot|\cdot)$  and  $\mathcal{E}_{i,t}(\cdot)$ . We distinguish between the first and the second conditional expectation as the model will be correctly specified only if the econometrician employs the right DGP. Moreover,  $T$  denotes the length of both the aggregate and average survey data, and  $L$  the number of agents. We denote by  $\theta \in \Theta$  the parameters of interest with  $\Theta_\alpha^{CS}$  the corresponding  $1 - \alpha$  level confidence set, and by  $q_j(\cdot; \theta)$  a measurable function. Bold capital letters e.g.  $\mathbf{Y}_t$  denote a vector of length  $\tau$  containing  $\{Y_j\}_{j \leq \tau}$ . The operator  $\rightarrow_p$  signifies convergence in probability and the operator  $\rightarrow_d$  convergence in distribution;  $\mathcal{N}(\cdot, \cdot)$  is the Normal distribution whose cumulative is  $\Phi(\cdot, \cdot)$ ;  $\|\cdot\|$  is the Euclidean norm unless otherwise stated;  $\perp$  signifies the orthogonal complement and  $\emptyset$  the empty set. We denote by  $(\Omega, \mathcal{S}_y, \mathbb{P})$  the probability triple for the observables to the econometrician, where  $\mathcal{S}_y = \sigma(Y)$ , is the sigma-field generated by  $Y$  and  $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t \geq 0}, \mathbb{P})$  the corresponding filtered probability space. Finally we denote by  $\odot$  and  $\oslash$  the Hadamard product and division respectively.

## 2. ROBUSTNESS IN PARTIAL EQUILIBRIUM

We first motivate our approach by illustrating the special case of observing a single household receiving a random exogenous endowment  $y_{i,t}$  and making consumption-savings decisions  $(c_{i,t}, s_{i,t})$  in a partial equilibrium context. As in [Zeldes \(1989\)](#), wealth  $w_{i,t}$  earns a riskless return  $R$  and there is a borrowing limit at zero:

$$\begin{aligned} & \max_{\{c_{i,t}\}_{t=1}^{\infty}} \mathbb{E}_0 \sum_{t=1}^{\infty} \beta^t \frac{c_{i,t}^{1-\omega} - 1}{1-\omega} \\ \text{s.t.} \quad & y_{i,t} = s_{i,t} + c_{i,t} \\ & w_{i,t+1} = R w_{i,t} + s_{i,t} \\ & w_{i,t+1} \geq 0 \end{aligned}$$

The corresponding consumption Euler equation is distorted by the non-negative Lagrange multiplier on the occasionally binding liquidity constraint, denoted by  $\lambda_{i,t}$ :

$$(1) \quad U'(c_{i,t}; \theta) = \beta \mathbb{E}_t(1 + r_{t+1})U'(c_{i,t+1}; \theta) + \lambda_{i,t}$$

Whether the borrowing constraint is binding depends on the household's expectations about future income, and therefore on its information set. Consequently, condition (1) is a conditional moment inequality, and for any conformable variable  $z_{i,t}$  that belongs to the household's information set, the following unconditional moment inequality holds:

$$(2) \quad \mathbb{E} [(U'(c_{i,t}; \theta) - \beta(1 + r_{t+1})U'(c_{i,t+1}; \theta)) \otimes z_{i,t}] \geq 0$$

Given the inequality, there is no unique vector  $(\theta, \beta)$  that satisfies it, and we therefore no longer have point identification. In order to derive explicit identification regions, we adopt the approximation of Hall (1978), which for CRRA utility and constant interest rate implies the following law of motion for consumption:

$$(3) \quad c_{i,t+1} = \mu c_{i,t} + \epsilon_{i,t+1} + \tilde{\lambda}_{i,t}$$

where  $\tilde{\lambda}_{i,t} \equiv -(U''(c_{i,t+1}))^{-1} \lambda_{i,t}$ . Since  $\mu$  may be exceed one if  $\beta R > 1$  and vice verca<sup>3</sup>, we re-express 3 in terms of consumption growth,  $\Delta c_{i,t+1} = \tilde{\mu} c_{i,t} + \epsilon_{t+1} + \tilde{\lambda}_{i,t}$ .

Since the sign of  $\mathbb{E}_t \tilde{\lambda}_t$  is still positive, the identified set for  $\mu_0$  is :

$$(4) \quad \mu_{ID,1} := \left( 0, 1 + \frac{\mathbb{E} z_{i,t} \Delta c_{i,t+1}}{\mathbb{E} z_{i,t} c_{i,t}} \right] = (0, 1 + \hat{\mu}]$$

As we can readily see from (4), the inherent unobservability of  $\tilde{\lambda}_{i,t}$  is the root of the loss of point identification of  $\mu_0$ . An interesting question is whether additional

<sup>3</sup>In fact,  $\mu$  is equal to  $(\beta(1+r))^{-\frac{u'(c)}{cu''(c)}} > 0$ , and it will be constant for CRRA utility.

data can be informative about, or even point identify,  $\mu_0$ . The answer in this case is affirmative.

Suppose that we observe the dichotomous response of the household over time to a survey question that asks whether the household is (or expects to be) financially constrained. An honest household will answer positively whenever  $\lambda_{i,t} > 0$ . What this implies is that the time series of responses  $\{\tilde{\chi}_{i,t}\}_{t \leq N}$  for  $\tilde{\chi}_{i,t} \equiv \mathbf{1}(\lambda_{i,t} > 0)$  can be used to estimate  $\mathbb{P}_t(\lambda_{i,t} > 0)$ . The conditional probability distribution of  $\Delta c_{i,t+1}$ ,  $\mathbb{P}_t(\Delta c_{i,t+1} < u)$  for any  $u \in \mathbb{R}$  is equal to

$$(5) \quad \mathbb{P}_t(\epsilon_{i,t+1} < u - \tilde{\mu}c_{i,t})\mathbb{P}_t(\lambda_{i,t} = 0) + \mathbb{P}_t(\epsilon_{i,t+1} < u - \tilde{\mu}c_{i,t} - \lambda_{i,t})\mathbb{P}_t(\lambda_{i,t} > 0)$$

For  $\epsilon_{t+1} \sim N(0, \sigma_\epsilon^2)$ , (5) simplifies further to,

$$(6) \quad \Phi_{0, \sigma_\epsilon^2}(u - \tilde{\mu}c_{i,t})\mathbb{P}_t(\lambda_{i,t} = 0) + \Phi_{0, \sigma_\epsilon^2}(u - \tilde{\mu}c_{i,t} - \lambda_{i,t})\mathbb{P}_t(\lambda_{i,t} > 0)$$

A key observation is that the distortion  $\lambda_{i,t}$  is positive only when the state variables  $(w_{i,t}, y_{i,t})$  lie in a particular area of their domain  $\mathcal{A}$  which is determined by an unknown endogenous threshold  $(w_i^*, y_i^*)$ . A *complete* model would pin down, analytically or numerically this domain. Observing  $\tilde{\chi}_{i,t}$  dispenses us with the need to characterize  $\mathcal{A}$ . Once the threshold condition is satisfied,  $\lambda_{i,t}$  will be a linear(ized) function of  $(w_{i,t}, y_{i,t})$ .

Moreover, it is also reasonable to assume that  $\lambda_{i,t}$  is a function of the unobserved exogenous shocks  $\epsilon_{t+1}$ . We can therefore substitute  $\lambda_{i,t}$  for  $\lambda_1 c_{i,t} + \lambda_2 \epsilon_{i,t+1}$ , in (6), which yields the following quantile restriction:

$$(7) \quad \begin{aligned} \mathbb{P}_t(\Delta c_{i,t+1} < u) &= \Phi_{0, \sigma_\epsilon^2}(u - \tilde{\mu}c_{i,t})\mathbb{P}_t(\lambda_{i,t} = 0) + \Phi_{0, \sigma_\epsilon^2(1+\lambda_2)^2}(u - \tilde{\mu}_{ols}^c c_{i,t})\mathbb{P}_t(\lambda_{i,t} > 0) \\ &\geq \Phi_{0, \sigma_\epsilon^2}(u - \tilde{\mu}c_{i,t})\mathbb{P}_t(\lambda_{i,t} = 0) + \Phi_{0, \sigma_{ols}^2}(u - \tilde{\mu}_{ols} c_{i,t})\mathbb{P}_t(\lambda_{i,t} > 0) \end{aligned}$$

where  $\tilde{\mu}_{ols}^c$  is equivalent to the least squares estimate in the case of borrowing constraints with probability one. The inequality is derived for a class of functions  $\mathbb{P}_t(\lambda_{i,t} > 0)$ .

How does this additional inequality refine the identified set in (4)? Denote by  $\mu_{ID,2}$  the set of  $\mu$  consistent with the unconditional moment inequality constructed using  $z_t$  and (7). The simplest way to show the refinement of (4) is to show that there exists a  $\mu \in \mu_{ID,1}$  that does not belong to  $\mu_{ID,2}$ . For simplicity, let  $\tilde{\mu}_0 = 0$ , which implies that  $\Delta c_{i,t+1} = \tilde{\chi}_{i,t}(\lambda_1 c_{i,t} + \lambda_2 \epsilon_{i,t+1}) + (1 - \tilde{\chi}_{i,t})\epsilon_{i,t+1}$ . This is equivalent to considering an economy in which  $\beta(1+r) < 1$  and the consumer has very high risk aversion ( $\omega \rightarrow \infty$ ). In Appendix A we show that using  $\mu_{ols}$  as the test point, (7) is not satisfied for  $p_t = 1(c_{i,t} < 0)$ , and therefore  $\mu_{ols} \notin \mu_{ID,2}$ . The upper bound in  $\mu_{ID}$  should thus be lower than  $\mu_{ols}$ . In words, as long as the agent has some positive unconditional probability of not being constrained (here  $\frac{1}{2}$ ), observing her responses provides additional information.

The identified set for the reduced form,  $\mu_{ID} = \mu_{ID,1} \cap \mu_{ID,2}$ , is informative in different ways. *First*, as already mentioned, given  $\mu_{ID}$ , we can recover the implied identified set for the risk aversion parameter,  $\omega_{ID}$ . As is typical in the empirical macro literature,  $(r, \beta)$  are not identified from the dynamic properties of the model but rather from other information, i.e steady states. In the case when  $\beta(1+r) = 1$ , risk aversion is unidentified,  $\omega_{ID} = \mathbb{R}^+$ , as consumption follows a random walk. For  $\beta(1+r) \neq 1$ ,

$$\omega_{ID} := \left\{ \omega \in \mathbb{R}^+ \cap \left[ \frac{\log(\beta(1+r))}{\log(1+\mu_{ID}^u)}, \frac{\log(\beta(1+r))}{\log(1+\mu_{ID}^l)} \right] \right\}$$

This interval has a very intuitive interpretation, and actually reflects restrictions on preferences implied by the presence of non-diversifiable income risk and observed behavior. For the sake of illustration let us focus on the identified set implied

by (4). If  $\beta(1+r) < 1$ , the household is impatient and does not accumulate wealth indefinitely. Using  $z_{i,t} = y_{i,t}$ , since income and consumption growth are negatively correlated, which is usually the case when liquidity constraints are present<sup>4</sup>,  $\omega_{ID,1} = \left\{ \omega \in \mathbb{R}^+ : \omega < \frac{\|\log(\beta(1+r))\|}{\log \mathbb{E}y_{i,t}c_{i,t} - \log(\mathbb{E}y_{i,t}c_{i,t} - \|\mathbb{E}y_{i,t}\Delta c_{i,t+1}\|)} \right\}$ . The stronger the negative correlation, the lower the upper bound on risk aversion, indicating the fact that the less risk averse household is not accumulating enough wealth to fully insure against income risk. On the contrary, weak negative correlation permits higher estimates of risk aversion. Similar to [Arellano, Hansen, and Sentana \(2012\)](#), the degree of under-identification therefore depends on this correlation, which is itself a function of  $(\omega, \beta, r)$ .

*Second*, the upper and lower bounds for  $\mu_{ID}$  can be used to do inference on other interesting quantities, including the model itself. For example, we can construct set estimates of distortions in consumption implied by liquidity constraints by plugging-in the identified set and averaging over the observations. In population, this produces an identified set for distortions:

$$(8) \quad \mathbb{E}\lambda_{i,t} \in \left[ \mathbb{E}c_{i,t+1} - \mu_{ID}^u c_{i,t}, \mathbb{E}c_{i,t+1} - \mu_{ID}^l c_{i,t} \right]$$

Another key observation is that although we have motivated the sign of  $\mathbb{E}_t\lambda_{i,t}$  by looking at the exogenous constraint of no borrowing, the sign is robust to different mechanisms, as long as they prevent the household from smoothing consumption. For example, a non zero constraint on next period wealth, i.e.  $w_{i,t+1} \geq \underline{b}$  or a constraint that is endogenous, i.e.  $w_{i,t+1} \geq \underline{b}(y_{i,t})$  where  $y_{i,t}$  is the relevant state variable, lead to an Euler equation which is distorted in the same direction. Therefore, the identified set of distortions will be a priori consistent with different mechanisms generating liquidity constraints.

---

<sup>4</sup>[Deaton \(1991\)](#)

A mechanism will not be rejected as long as it generates distortions that will statistically belong to the identified set for  $\mathbb{E}\lambda_t$ . Moreover, we will see that in general equilibrium, survey data will generate additional moment restrictions which can potentially add information. The identified set is expected to be refined, and this refinement can potentially provide power to reject alternative models that cannot be rejected using bound (4).

In the rest of the paper we will gradually characterize the case for general equilibrium models and what information aggregated surveys can provide, based on similar assumptions to those made in the simple example. To do so, we will provide below the benchmark general equilibrium model.

### 3. TOWARDS GENERAL EQUILIBRIUM

In the last section we considered liquidity constraints for a single household in a partial equilibrium context. In this section, we consider additional types of frictions by including capital and a representative firm in the economy and set up the general equilibrium framework. Building up from the previous section, consider a simple Real Business Cycle (RBC) model, in which dynastic households with Constant Relative Risk Aversion (CRRA) preferences form expectations about key state variables and make consumption - savings decisions. We denote by  $\Lambda_t(i)$  the distribution of agents at time  $t$ , and the corresponding aggregate variables by capital letters i.e.  $X_t \equiv \int x_{i,t} d\Lambda_t(i)$ . Households rent capital ( $k_{i,t}$ ) to a representative firm, which is used for production, and receive a share ( $\eta_{i,t}$ ) of profits made by the firm. Profits are simply the production of output using capital intensive technology with random productivity ( $Z_t K_t^\alpha$ ) minus the rents paid to all households ( $R_t K_t$ ), that is,  $pr_{i,t} = \eta_{i,t}(Z_t K_t^\alpha - R_t K_t)$ . Household income is therefore  $y_{i,t} = R_t k_{i,t} + pr_{i,t}$ . Individual investment decisions ( $l_{i,t}$ ) increase the availability of capital for next period, up to a certain level

of depreciation. The household's optimization problem is therefore as follows:

$$\begin{aligned} & \max_{\{c_{i,t}, k_{i,t+1}, w_{i,t+1}, l_{i,t}\}_1^\infty} \mathcal{E}_{i,0} \sum_{t=1}^{\infty} \beta^t \frac{c_{i,t}^{1-\omega} - 1}{1-\omega} \\ \text{s.t. } & w_{i,t+1} = R_t w_{i,t} + y_{i,t} - c_{i,t} - l_{i,t} \\ & k_{i,t+1} = (1-\delta)k_{i,t} + l_{i,t} \end{aligned}$$

Note that individual expectations  $\mathcal{E}_{i,0}$  are not necessarily formed with respect to the objective probability measure, nor with respect to the same information set. Denote by  $\xi_t$  the aggregation residual, and by  $\tilde{X}$  the percentage deviations of any aggregate variable  $X_t$  from aggregate steady state  $x_{ss}$ . Then the system of aggregate linearized<sup>5</sup> equilibrium and market clearing conditions are as follows:

$$\begin{aligned} -\omega \tilde{C}_t - \xi_{C,t} - \beta r_{ss} \bar{\mathcal{E}}_t \tilde{R}_{t+1} + \omega \bar{\mathcal{E}}_t \tilde{C}_{t+1} &= 0 \\ y_{ss} \tilde{Y}_t - c_{ss} \tilde{C}_t - i_{ss} \tilde{I}_t &= 0 \\ \tilde{K}_{t+1} - (1-\delta) \tilde{K}_t - \delta \tilde{I}_t &= 0 \\ \tilde{Y}_t - \tilde{Z}_t - \alpha \tilde{K}_t - \xi_{Y,t} &= 0 \\ \tilde{R}_t - \tilde{Z}_t + (1-\alpha) \tilde{K}_t - \xi_{R,t} &= 0 \end{aligned}$$

Under approximate linearity, that is, when  $\xi_t$  is negligible<sup>6</sup>, we can obtain the aggregate decision rules for  $X_t$ , which will depend on predetermined capital  $\tilde{k}_{i,t}$  and on aggregate conditional expectations on productivity for some  $j$  periods ahead,  $\bar{\mathcal{E}}_t \tilde{Z}_{t+j}$ .

<sup>5</sup>Note that here we follow the linear approximation to the first order conditions as widely used in the DSGE literature, while in the last section we followed Hall (1978), where the variables are not in percentage deviation from steady state. The analysis is nevertheless largely unaffected.

<sup>6</sup>Absence of frictions is one of the reasons for which approximate linearity holds.

In the previous section, we focused on the consumption decision of agents by taking the interest rate as exogenous and fixed, that is  $R_t = R$ . In this section,  $R_t$  is endogenous as it is determined by the marginal productivity of capital, and is subject to aggregate risk. With regard to the aggregate implications of the borrowing constraints introduced in section 2, the characterization of individual behavior is identical. In general equilibrium, what needs to be taken into account is the determination of  $R_t$ . We will formally deal with general equilibrium in the next section. Moreover, in this section the idiosyncratic risk to income is  $\eta_{i,t}$  while there is also aggregate risk,  $Z_t$ . Compared to the last section, the introduction of aggregate risk does not alter the observable implications on individual consumption behavior. This is for the reason that although aggregate risk complicates forecasting for households as predicting  $R_t$  requires tracking  $\Lambda_t(i)$ <sup>7</sup>, the inequality in (2) is valid irrespective of the presence of aggregate uncertainty.

In what follows we will analyze the equilibrium law of motion for investment, where we can introduce additional types of frictions, in particular those that arise from the production side of the economy. As in the case of consumption, we first analyze the aggregate investment decision rule in the frictionless case, which should satisfy the following second order difference equation:

$$(9) \quad \frac{\bar{s}_0}{\alpha} \tilde{\mathcal{E}}_t \tilde{I}_{t+1} - \left( 1 + \frac{(1-\alpha)(1-\bar{s}_0) + \bar{s}_0 \omega}{\alpha \omega} \right) \tilde{I}_t + \tilde{I}_{t-1} = \frac{1}{\alpha} \tilde{Z}_t$$

where  $\bar{s}_0 \equiv \frac{I_{ss,0}}{Y_{ss,0}}$ , the frictionless savings ratio.

A sufficient condition for a saddle point solution for the latter is  $\alpha > \bar{s}$ , that is, the steady state savings rate should be smaller than the elasticity of output with respect to capital. Denoting by  $\rho_{1,0}, \rho_{2,0}$  the two roots of the corresponding lag polynomial and letting  $\rho_{2,0}$  be the unstable root, the aggregate frictionless investment has the

---

<sup>7</sup>Krusell and Anthony A. Smith (1998)

following law of motion:

$$(10) \quad \tilde{I}_t = \rho_{1,0}\tilde{K}_t + \frac{1}{\rho_{2,0}\bar{s}_0} \sum_{j=1}^{\infty} \bar{\mathcal{E}}_t \rho_{2,0}^{-j} \tilde{Z}_{t+j}$$

Consider now the case in which the true model has no real or nominal frictions and agents have rational expectations ( $\mathcal{P}_{t,i} = \mathbb{P}_{t,i}$ ) where the corresponding information set is  $\mathcal{F}_i \equiv (\eta_{i,t}, k_{i,t}, Z_t)$ . Following the literature, we assume that only a subset of  $\{\mathcal{F}_i\}_{i \leq N}$  is observed by the econometrician, i.e.  $\mathcal{F}_t = K_t$ . The decision rule used by the econometrician can then be rewritten as:

$$(11) \quad \tilde{I}_t = \rho_{1,0}\tilde{K}_t + \frac{1}{\rho_{2,0}\bar{s}_0} \sum_{j=0}^{\infty} \rho_{2,0}^{-j} \mathbb{E}_t(\tilde{Z}_{t+j} | \sigma(\mathcal{F}_t)) + \frac{1}{\rho_{2,0}\bar{s}_0} e_t$$

where

$$\begin{aligned} e_t &= \sum_{j=0}^{\infty} \rho_{2,0}^{-j} \int \mathbb{E}(\tilde{Z}_{t+j} | \mathcal{F}_{i,t}) d\Lambda_t(i) - \sum_{j=0}^{\infty} \rho_{2,0}^{-j} \mathbb{E}(\tilde{Z}_{t+j} | \mathcal{F}_t) \\ &= \sum_{j=0}^{\infty} \rho_{2,0}^{-j} \mathbb{E}(\tilde{Z}_{t+j} | \tilde{Z}_t, \tilde{K}_t) - \sum_{j=0}^{\infty} \rho_{2,0}^{-j} \mathbb{E}(\tilde{Z}_{t+j} | \tilde{K}_t) \end{aligned}$$

The second equality holds due to the linearity of conditional expectations and aggregation. Then, for any  $\tilde{K}_{t-j, j \geq 0}$ -measurable function  $\phi(\cdot)$ , and using the law of total expectations, the following moment equality holds<sup>8</sup>:

$$\mathbb{E} \tilde{K}_{t-j} e_t = 0$$

The case of interest in this paper is therefore when  $\int \mathcal{E}(\cdot | \mathcal{F}_{i,t}) d\Lambda_t(i) \neq \mathbb{E}^{M_f}(\cdot | \mathcal{F}_t)$ , where  $M_f$  is the frictionless probability model- or the benchmark model in the general case<sup>9</sup>. Here the agents and the econometrician not only have different information, but they also have a different structure of the economy in mind.  $\mathbb{E}^{M_f}(\cdot | X_t)$  is consistent

<sup>8</sup>Note that we have not used any knowledge for the exogeneity of  $\tilde{Z}_t$  for this result to hold i.e. it would hold trivially in the case of an exogenous productivity shock.

<sup>9</sup>To avoid confusion, I use frictionless and benchmark model interchangeably.

with frictionless behavior and we thus implicitly assume that the frictionless part of the model is well specified. The mismatch between the agents' expectations and the econometrician's prediction could be due to differences in the models and/or information sets, which can lead to different decision rules, and these differences do not vanish on average.

Therefore, denoting aggregate optimal investment in the presence of frictions by  $\tilde{I}_{con,t}$ , which is the data generating process, and by  $I_t^*$  the frictionless investment rule used by the econometrician, we can represent these differences in terms of the econometrician's observables as follows:

$$\begin{aligned}\tilde{I}_{con,t} &= \rho_{1,0}\tilde{K}_t + \frac{1}{\rho_{2,0}\bar{s}} \sum_{j=0}^{\infty} \rho_{2,0}^{-j} \mathbb{E}_t(\tilde{Z}_{t+j} | \sigma(\mathcal{F}_t)) + \frac{1}{\rho_{2,0}\bar{s}} e_t + \tilde{\lambda}_t \\ \tilde{\lambda}_t &= \tilde{I}_{con,t} - I_t^*\end{aligned}$$

or equivalently, as

$$(12) \quad \mathbb{E} \left( \tilde{I}_{con,t} - \rho_{1,0}\tilde{K}_t - \frac{1}{\rho_{2,0}\bar{s}} \sum_{j=0}^{\infty} \rho_{2,0}^{-j} \tilde{Z}_{t+j} - e_t - \tilde{\lambda}_t | \sigma(\mathcal{F}_t) \right) = \mathbb{E}(\tilde{\lambda}_t | \sigma(\mathcal{F}_t))$$

Similar to the case of consumption with liquidity constraints, real, nominal or informational frictions therefore generate a “wedge”  $\tilde{\lambda}_t$  with non zero conditional mean and therefore a moment inequality. We will consider a general example of constraints in adjusting capital, which can be rationalized by ad hoc adjustment costs in capital accumulation, occasionally binding constraints i.e. capital irreversibility and financial frictions. The specific examples together with an example on non-rational expectations can be found in the Appendix B (online). Below, we present the general example.

**3.0.1. Capital Adjustment Constraints.** Assuming full capital depreciation, [Wang and Wen \(2012\)](#) show that collateral constraints can be represented by an aggregate

function  $\psi(\iota_t)$  of the investment rate,  $\iota_t \equiv \frac{I_t}{K_t}$ . Note that this representation can be rationalized with different setups, including those in which there is a distribution of investment efficiency shocks across firms. Moreover, such setups can be further microfounded by different mechanisms i.e. financial frictions à la [Bernanke, Gertler, and Gilchrist \(1999\)](#). Focusing on exogenous collateral constraints, i.e. the borrowing limit does not depend on the asset value of the collateral<sup>10</sup>, they imply the following subsystem of linearized aggregate equilibrium conditions:

$$(13) \quad \tilde{K}_{t+1} = \psi'_{ss} \tilde{I}_t + (1 - \psi'_{ss}) \tilde{K}_t$$

$$(14) \quad \tilde{Q}_t = -\psi'^{-1}_{ss} \psi''_{ss} (\tilde{I}_t - \tilde{K}_t)$$

$$(15) \quad \tilde{Q}_t = \omega \tilde{C}_t - \omega \mathbb{E}_t C_{t+1} + \Psi_{ss,1} \mathbb{E}_t \tilde{R}_{t+1} + \mathbb{E}_t \tilde{Q}_{t+1} + (\psi'_{ss} - 1) \mathbb{E}_t (\tilde{I}_{t+1} - \tilde{K}_{t+1})$$

where  $\Psi_{ss,1} \equiv 1 + \beta(\psi'_{ss} - 1)$ .

Denoting by  $\Psi_{ss,2} \equiv \omega(\psi'_{ss} \bar{s}(\alpha - \bar{s}) + \bar{s}) - \bar{s}(1 - \bar{s})\psi'^{-1}_{ss} \psi''_{ss}$ ,  $\gamma_1 \equiv \Psi_{ss,2}^{-1}(\omega \bar{s} - \psi'^{-1}_{ss} \psi''_{ss} \bar{s}(1 - \bar{s}))$ ,  $\gamma_2 \equiv \Psi_{ss,2}^{-1} \omega$  and  $\gamma_3 \equiv \Psi_{ss,2}^{-1} \Psi_{ss,1}(1 - \bar{s})$  the first order condition (9) becomes

$$(16) \quad \gamma_1 \mathbb{E}_t \tilde{I}_{t+1} - (1 + \gamma_1 + (1 - \alpha)\gamma_3) \tilde{I}_t + \tilde{I}_{t-1} = \gamma_2 \tilde{Z}_t$$

whose solution generalizes to:

$$\tilde{I}_{con,t} = \rho_1(\gamma) \tilde{K}_t + \frac{1}{\rho_2(\gamma) \bar{s}(\gamma)} \tilde{Z}_t + \frac{1}{\rho_2} e_t + \tilde{\lambda}_t$$

For  $\rho_i(\gamma_0) \equiv \rho_{i,0}$ ,  $\bar{s}(\gamma_0) \equiv \bar{s}_0$ , the corresponding distortion due to capital adjustment constraints is equivalent to

$$\tilde{\lambda}_t = (\rho_1(\gamma) - \rho_1(\gamma_0)) \tilde{K}_t + \left( \frac{1}{\rho_2(\gamma) \bar{s}(\gamma)} - \frac{1}{\rho_2(\gamma_0) \bar{s}(\gamma_0)} \right) \tilde{Z}_t + \left( \frac{1}{\rho_2(\gamma)} - \frac{1}{\rho_2(\gamma_0)} \right) e_t$$

<sup>10</sup>Endogenous collateral constraints have a different representation than exogenous constraints. Nevertheless, as [Wang and Wen \(2012\)](#) note, endogenous collateral constraints imply a form of aggregate investment externality which implies insufficient level of investment relative to the case of exogenous constraints. This simply reinforces the distortions we derive based on exogenous constraints.

In order to determine the sign of  $\mathbb{E}(\tilde{\lambda}_t|\sigma(\mathcal{F}_t))$ , it is sufficient to look at the sign of the coefficient of  $\tilde{K}_t$ . Using a mean value expansion around zero, the conditional mean is  $\mathbb{E}(\tilde{\lambda}_t|\sigma(\mathcal{F}_t)) = \rho'_1(\bar{\gamma})\tilde{K}_t$  where  $\rho'_1(\bar{\gamma}) \equiv \underset{(-)}{\rho_1(\bar{\gamma})'(\gamma_1 - \gamma_{1,0})} + \underset{(+)}{\rho'_1(\bar{\gamma})(\gamma_3 - \gamma_{3,0})} \underset{(-)}{\leq} 0$ .

Frictions in capital accumulation therefore generate negative endogenous distortions, which implies that  $\mathbb{E}\tilde{\lambda}_t\tilde{K}_t \leq 0$ <sup>11</sup>.

The preceding example dealt with distortions in the decision rules, which directly affect- and are therefore informative for- the transmission of shocks and welfare. Moreover, they can be easily linked to additional data. Nevertheless, frictions are always justified using the first order necessary conditions for agents' optimal decisions. In the next section, we present the general case together with a representation result, which proves useful in translating distortions to the first order conditions into observationally equivalent general equilibrium distortions to decision rules. One can therefore directly use the latter. Because we work with linearized models, second or higher order effects will be ignored. However, moment inequalities, would also appear in nonlinear models. The way we treat frictions does not depend on assumptions regarding the approximation error.

#### 4. PERTURBING THE FRICTIONLESS MODEL

The general framework involves a system of expectational equations. Denote by  $x_{i,t}$  the endogenous individual state, by  $z_{i,t}$  the exogenous individual state, and by  $X_t = \int x_{i,t}d\Lambda(i)$  and  $Z_t = \int z_{i,t}d\Lambda(i)$  the corresponding aggregate states. The optimality

<sup>11</sup>The signs hold uniformly over  $\gamma$ , for any twice differentiable function  $\psi$ . This is straightforward for the derivatives, while for  $\gamma_i - \gamma_{i,0}$ , this gives rise to a quadratic inequality in  $\psi'_{ss}$  which holds for all plausible  $\psi'_{ss} \in (0, 1]$ .

conditions characterizing the individual decisions are as follows:

$$(17) \quad \begin{aligned} G(\theta)x_{i,t} &= F(\theta)\mathcal{E}_{i,t} \left( \left( \begin{array}{c} x_{i,t+1} \\ X_{t+1} \end{array} \right) \middle| x_{i,t}, z_{i,t}, X_t, Z_t \right) + L(\theta)z_{i,t} \\ z_{i,t} &= R(\theta)z_{i,t-1} + \epsilon_{i,t} \end{aligned}$$

where  $\mathbb{E}(\epsilon_{i,t}) = 0$ . We assume that the coefficients of the behavioural equations are common across agents. Relaxing this assumption would make the notation more complicated, but would not change the essence of our argument. We could also specify equilibrium conditions that involve past endogenous variables but this is unnecessary as we can always define dummy variables of the form  $\tilde{x}_{i,t} \equiv x_{i,t-1}$  and enlarge the vector of endogenous variables to include  $\tilde{x}_{i,t}$ . Aggregating across individuals, the economy can be characterized by the following system:

$$(18) \quad G(\theta)X_t = F(\theta) \int \mathcal{E}_{i,t} \left( \left( \begin{array}{c} x_{i,t+1} \\ X_{t+1} \end{array} \right) \middle| x_{i,t}, z_{i,t}, X_t, Z_t \right) d\Lambda_t(i) + L(\theta)Z_t$$

$$(19) \quad Z_t = R(\theta)Z_{t-1} + \epsilon_t$$

We will refer to the economy with frictions as the triple  $(H(\theta), \Lambda, \mathcal{E}_i)$  where

$$H(\theta) \equiv (\text{vec}(G(\theta))^T, \text{vec}(F(\theta))^T, \text{vec}(L(\theta))^T, \text{vec}(R(\theta))^T, \text{vech}(\Sigma_\epsilon)^T).$$

We partition the vector  $\theta$  into two subsets,  $(\theta_1, \theta_2)$  where  $\theta_2$  collects the parameters characterizing the presence and intensity of frictions. Thus, setting  $\theta_2 = 0$ , reduces the model to the frictionless economy and this is without loss of generality. Furthermore, in an economy with no frictions, prices efficiently aggregate all the information. Thus

there is no need to distinguish between individual and aggregate information when predicting aggregate state variables.

For  $w_{i,t} \equiv (x_{i,t}, X_{i,t}, z_{i,t}, Z_{i,t})$ , when agents are rational, and the model is linear ( or linearized) aggregate expectations for  $(x_{i,t+1}^T, X_{t+1}^T)^T$  become as follows:

$$\begin{aligned}\mathbb{E}x_{i,t+1} &= \iint x_{i,t+1} p_i(x_{i,t+1}, X_{t+1} | w_{i,t}) d(x_{i,t+1}, X_{t+1}) d\Lambda_t(i) \\ &= \iint x_{i,t+1} p(x_{i,t+1} | w_{i,t}) d(x_{i,t+1}) d\Lambda_t(i) \\ &= \int P_{i,1} x_{i,t} d\Lambda(i) + P_2 X_t + \int P_{i,3} z_{i,t} d\Lambda_t(i) + P_4 Z_t\end{aligned}$$

$$\begin{aligned}\mathbb{E}X_{t+1} &= \iint X_{t+1} p_i(x_{i,t+1}, X_{t+1} | w_{i,t}) d(x_{i,t+1}, X_{t+1}) d\Lambda_t(i) \\ &= \int X_{t+1} p(X_{t+1} | X_t, Z_t) d(X_{t+1}) \\ &= P_5 X_t + P_6 Z_t \equiv \mathbb{E}(X_{t+1} | X_t, Z_t)\end{aligned}$$

where  $P_{j,j=1..6}$  are the coefficients of the linear projection. By Rational expectations and since the coefficients  $(G, F, L)$  are common across  $i$ , equilibrium consistency requires  $P_{i,1} = P_1, P_{i,2} = P_2$  and therefore  $P_1 + P_2 = P_5$  and  $P_3 + P_4 = P_6$ . Thus, as expected, aggregate conditional expectations collapse to  $\mathbb{E}(X_{t+1} | X_t, Z_t)$ , and the frictionless economy,  $(H(\theta_1, 0), \Lambda, \mathbb{E})$ , can be summarized by the equilibrium conditions:

$$(20) \quad G(\theta_1, 0)X_t = F(\theta_1, 0)\mathbb{E}_t(X_{t+1} | X_t, Z_t) + L(\theta_1, 0)Z_t$$

$$(21) \quad Z_t = R(\theta)Z_{t-1} + \epsilon_t$$

Using the decision rule in the expectational system (20) and solving for the undetermined coefficients <sup>12</sup>, a Rational Expectations equilibrium for  $(H(\theta_1, 0), \Lambda, \mathbb{E})$  holds under the following conditions:

. **ASSUMPTION-EQ**

There exist unique matrices  $P^*(\theta_1, 0)$ ,  $Q^*(\theta_1, 0)$  satisfying:

$$(F(\theta_1, 0)P^*(\theta_1, 0) - G^*(\theta_1, 0))P^*(\theta_1, 0) = 0$$

$$(R(\theta)^T \otimes F(\theta_1, 0) + I_z \otimes (-F(\theta_1, 0)P^*(\theta_1, 0) + G^*(\theta_1, 0)))\text{vec}(Q(\theta_1, 0)) = -\text{vec}(L(\theta_1, 0))$$

such that  $X_t = P^*(\theta_1, 0)X_{t-1} + Q^*(\theta_1, 0)Z_t$  is a competitive equilibrium.

Since we know the model up to  $\theta_2 = 0$ , we rearrange the equations of the economy with frictions into the known and the unknown part of the specification. Adding and subtracting the first order conditions of the frictionless economy we get that:

$$(22) \quad G(\theta_1, 0)X_t = F(\theta_1, 0)\mathbb{E}_t(X_{t+1}|X_t) + L(\theta_1, 0)Z_t + \mu_t$$

$$\begin{aligned} \mu_t &\equiv -(G(\theta) - G(\theta_1, 0))X_t + (L(\theta) - L(\theta_1, 0))Z_t \\ &\quad + (F(\theta) - F(\theta_1, 0)) \int \mathcal{E}_{it} \left( \left( \begin{array}{c} x_{i,t+1} \\ X_{t+1} \end{array} \right) | x_{i,t}, X_t \right) d\Lambda_t(i) \\ &\quad + F(\theta_1, 0) \left( \int \mathcal{E}_{it} \left( \left( \begin{array}{c} x_{i,t+1} \\ X_{t+1} \end{array} \right) | x_{i,t}, X_t \right) d\Lambda_t(i) - \mathbb{E}_t(X_{t+1}|X_t) \right) \end{aligned}$$

This system of equations cannot be solved without knowing  $\mu_t$ . Nevertheless, we characterize the relationship between  $\mu_t$  and a set of candidate decision rules that depend on the endogenous states and some unobserved process,  $\lambda_t$ . Proposition 1 states sufficient conditions such that decision rules are consistent with  $\mu_t$ .

<sup>12</sup>See for example [Marimon and Scott \(1998\)](#).

**Proposition 1.** *Given:*

- (1) *The perturbed system of equilibrium conditions (22) where  $\mathbb{E}_t \mu_t \geq 0$*
- (2) *A distorted aggregate decision rule  $X_t^* = X_t^{f,RE} + \lambda_t$  where  $X_t^{f,RE} = \int x_{i,t}^{RE} d\Lambda(i)$  is the Rational Expectations equilibrium of  $(H(\theta_1, 0), \Lambda, \mathbb{E})$  and*
- (3) *A  $\lambda_t$  vector process such that  $\lambda_t = \lambda_{t-1}\Gamma + \nu_t$  for some real-valued  $\Gamma$ :*

*If there exists a non-empty subset of  $\Theta_1$  that satisfies*

$$(23) \quad \mathbb{E}_t(F(\theta_1, 0)\Gamma - G(\theta_1, 0))\lambda_t = -\mathbb{E}_t\mu_t$$

*The following condition is satisfied for almost all subsets of  $\sigma(Y_{t-1})$  :*

$$\mathbb{E}_t(G(\theta_1, 0)X_t^* - F(\theta_1, 0)(\theta)X_{t+1}^* - L(\theta_1, 0)Z_t) \geq 0$$

*Proof.* See Appendix □

Proposition 1 states that if for an admissible parameter vector  $\theta_1 \in \Theta_1$  condition (24) is true, the decision rule  $X_t^* = X_t^{f,RE} + \lambda_t$  generates the same restrictions as those implied using the perturbed (by  $\mu_t$ ) first order equilibrium conditions. We focus on parameter vectors that yield determinate and stable equilibria in the frictionless economy, which implies a restriction on the stochastic behavior of  $\lambda_t$ .

While it is true that in many cases the sign of  $\mathbb{E}_t\lambda_t$  can be directly deduced from the distortions to equilibrium conditions, Proposition 1 can also be useful in practice, as we can deduce the sign of the distortions to decision rules without solving for the otherwise unknown expectational system using (23)<sup>13</sup>. Moreover, Proposition 1 states conditions under which we can guarantee that the moment inequality restrictions are

---

<sup>13</sup>In more complicated cases where knowledge of  $\Gamma$  is strictly required, condition (23) can always be checked ex post.

going to be consistent with the rest of the model, and therefore the implied reduced form.

## 2. Analytical example for Capital Adjustment constraints

As illustrated in the previous section, distortions arise in the capital accumulation equation, Tobin's Q and the Euler equation. Reducing the system to its minimal representation in terms of capital and investment i.e.  $X_t \equiv (\tilde{I}_t, \tilde{K}_{t+1})$ , we have that:

$$\begin{bmatrix} 1 & \gamma_{1,0} + (1 - \alpha)\gamma_{3,0} \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \tilde{I}_t \\ \tilde{K}_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbb{E}_t \tilde{I}_{t+1} \\ \mathbb{E}_t \tilde{K}_{t+2} \end{bmatrix} + \begin{bmatrix} \gamma_{2,0} \\ 0 \end{bmatrix} \tilde{Z}_t + \begin{bmatrix} -\mu_{I,t} \\ 0 \end{bmatrix}$$

where the Investment Euler equation has a positive distortion as the marginal cost of deferring consumption increases. We can then derive  $(\tilde{\lambda}_{I,t}, \tilde{\lambda}_{K,t+1})$ , and consequently distortions to all other endogenous variables i.e. consumption. Since  $\tilde{K}_{t+1} = \tilde{I}_t$ ,  $\Gamma_{11} = \Gamma_{22} < 1$ . For  $\Omega := 1 - \Gamma_{11} + \gamma_{1,0} + (1 - \alpha)\gamma_{3,0}$ , computing  $\mathbb{E}_t(F(\theta_1, 0)\Gamma - G(\theta_1, 0))\tilde{\lambda}_t + \tilde{\mu}_t = 0$  for the reduced system and solving for the equilibrium distortions, leads to the following result:

$$\begin{bmatrix} \tilde{\lambda}_{I,t} \\ \tilde{\lambda}_{K,t+1} \end{bmatrix} = -\frac{1}{\Omega} \begin{bmatrix} \tilde{\mu}_{I,t} \\ \tilde{\mu}_{I,t} \end{bmatrix}$$

Although trivial, this example illustrates how cross equation restrictions determine the sign of all distortions, can then be utilized to construct moment inequalities.

It is useful to note that in certain situations one may be interested in characterizing frictions over time (and not just on average), and therefore we need to obtain a conditional model that generates  $\lambda_t$ . This requires imposing more restrictions on the stochastic behavior of  $\lambda_t$ . In the supplementary material, we adopt a model uncertainty approach in the spirit of [Hansen and Sargent \(2005\)](#); [Hansen \(2013\)](#) and we show that restricting the class of distributions for aggregate shocks is enough to obtain a unique conditional model for  $\lambda_t$ . For the rest of this paper we focus on the identification using a set of unconditional moment restrictions.

## 5. IDENTIFICATION AND ESTIMATION

In this section we provide a formal treatment of identification in a linear(ized) Dynamic Stochastic General Equilibrium (DSGE) model based on moment inequality restrictions. First, we illustrate how our statistical representation of a DSGE model relates to the state space representation that is typically used for estimation. Building on [Komunjer and Ng \(2011\)](#), we will show necessary and sufficient conditions for partial identification of the model arising from the theoretical moment inequalities. We also show conditions under which *conditionally* over-identifying inequalities provide a more informative (smaller) identified set.

We base the analysis on the innovation representation of the solution to the linear(ized) DSGE model. This is the natural representation to use when there are differences in information between economic agents and the econometrician, as it takes into account that not all the state variables relevant the decision of agents are observable. We consider the following class of models:

$$\begin{aligned}\hat{X}_{t+1|t} &= \underset{n_X \times n_X}{A(\theta)} \hat{X}_{t|t-1} + \underset{n_X \times n_X}{K_t(\theta)} a_t \\ Y_t^o &= \underset{n_Y \times n_x}{C(\theta)} \hat{X}_{t|t-1} + a_t\end{aligned}$$

where  $K_t(\theta)$  is the Kalman gain and  $a_t$  is the one-step ahead forecast error which could be derived from the corresponding state space representation:

$$\begin{aligned}X_{t+1} &= \underset{n_X \times n_X}{A(\theta)} X_t + \underset{n_X \times n_X}{B(\theta)} \epsilon_{t+1} \\ Y_t &= \underset{n_Y \times n_X}{C(\theta)} X_t + \lambda_t = \underset{n_Y \times n_X n_X \times n_X}{C(\theta) A(\theta)} X_{t-1} + \underset{n_Y \times n_X n_X \times n_X}{C(\theta) B(\theta)} \epsilon_t + \lambda_t \\ &= \underset{n_Y \times n_X}{\tilde{C}(\theta)} X_{t-1} + \underset{n_Y \times n_X}{D(\theta)} \epsilon_t + \lambda_t\end{aligned}$$

where  $\epsilon_t$  is the innovation to the shock vector  $Z_t$ .

By construction,

$$a_t = \lambda_t + C(\theta)A(\theta)(X_{t-1} - \hat{X}_{t-1|t-1}) + C(\theta)B(\theta)\epsilon_t$$

Therefore, the forecast error is a combination of the true aggregate innovations to the information sets of the agents,  $\epsilon_t$ , the estimation error of the state variable,  $X_{t-1} - \hat{X}_{t-1,t-1}$ , and the frictions,  $\lambda_t$ . Let  $N(\theta) \equiv \text{vec}(A(\theta)', B(\theta)', C(\theta)')$ , and assume that  $\mathbb{E}(\epsilon_t | \sigma(\mathcal{F}_t)) = 0$ , and  $\mathbb{E}(\epsilon_t \epsilon_s' | \sigma(\mathcal{F}_t)) = \mathbf{1}(s = t) \Sigma_{\epsilon_t}$ , where  $\Sigma_{\epsilon_t} > 0$ . Given  $\mathbb{E}(\lambda_t | \sigma(\mathcal{F}_{t-1})) \geq 0$ , we define the following conditional moment restriction:

$$(24) \quad \mathbb{E}(Y_t^o - C(\theta) \hat{X}_{t|t-1} | \sigma(\mathcal{F}_{t-1})) \geq 0$$

$n_Y \times n_x$

For any inequality preserving function  $\phi(\cdot)$  of a random vector  $\mathbf{Y}_{t-1}$  that belongs to the information set of the econometrician, the following holds:

$$\mathbb{E}(Y_t^o - C(\theta) \hat{X}_{t|t-1}) \phi(\mathbf{Y}_{t-1}) = \mathbb{E} \mathcal{V}(\mathbf{Y}_{t-1}) \phi(\mathbf{Y}_{t-1}) \geq 0$$

$n_Y \times n_x$

for a random function  $\mathcal{V}(\mathbf{Y}_{t-1}) \in [0, \infty]$ .

In order to study identification through estimating equations we need to make assumptions about the local identification of  $\Theta_0$ , given the value of  $\mathbb{E} \mathcal{V}(\mathbf{Y}_{t-1}) \phi(\mathbf{Y}_{t-1})$ . We resort to sufficient conditions that make the mapping from  $\theta$  to the solution of the model regular, and thus assume away population identification problems (see, for example [Canova and Sala \(2009\)](#)). We assume that  $\Theta$  belongs to a compact subset of  $\mathbb{R}^{n_\theta}$ . Since certain parameters are naturally restricted, e.g. discount factors, persistence parameters or fractions of the population, and others cannot take excessively high or low values, assuming compactness is innocuous. We also need to acknowledge that due to *cross - equation restrictions*, which we denote by  $\mathfrak{L}(\theta) = 0$ , the number of observables used in the estimation need not be equal to the cardinality of  $\Theta$ , i.e.,  $n_y < n_\theta$ . [Komunjer and Ng \(2011\)](#) provide the necessary and sufficient conditions

for local identification of the DSGE model from the auto-covariances of the data. We reproduce them below, with the minor modification that Assumption **LCI-6** holds for any element of the identified set  $\Theta_0$ .

. **ASSUMPTION -LCI (Local Conditional Identification)**

- (1)  $\Theta$  is compact and connected
- (2) (Stability) For any  $\theta \in \Theta$  and for any  $z \in \mathbb{C}$ ,  $\det(zI_{n_x} - A(\theta)) = 0$ , implies  $|z| < 1$
- (3) For any  $\theta \in \Theta$ ,  $D(\theta)\Sigma_e D(\theta)'$  is non-singular
- (4) For any  $\theta \in \Theta$ , (i) The matrix  $(K(\theta) A(\theta)K(\theta) \dots, A(\theta)^{n_x-1}K(\theta))$  has full row rank and  $(C(\theta)' A(\theta)'C(\theta)' \dots, A(\theta)^{n_x-1}C(\theta)')$  has full column rank.
- (5) For any  $\theta \in \Theta$ , the mapping  $N : \theta \mapsto N(\theta)$  is continuously differentiable
- (6) Rank of  $\Delta^{NS}(\theta)$ <sup>14</sup> is constant in a neighborhood of  $\theta_0 \in \Theta_I$  and is equal to  $n_\theta + n_x^2$

**Lemma 3.** Given  $\mathcal{V}_i(\cdot) \in [\underline{\mathcal{V}}(\cdot), \bar{\mathcal{V}}(\cdot)]$ , and Assumption **LCI**,  $\theta$  is locally conditionally identified at any  $\theta_0$  in  $\Theta_I$  from the auto-covariances of  $Y_t$ .

*Proof.* See Appendix □

Given the maintained assumptions, we next characterize the identified set implied by the restrictions using macroeconomic data.

<sup>14</sup> For  $\delta^{NS}(\theta, T) = (\text{vec}(TA(\theta)T^{-1})^T, \text{vec}(TK(\theta))^T, \text{vec}(C(\theta)T^{-1})^T, \text{vech}(\Sigma_\alpha(\theta))^T)^T$ ,  $\Delta^{NS}(\theta_0) = (\frac{\partial \delta^{NS}(\theta, I_{n_x})}{\partial \theta}, \frac{\partial \delta^{NS}(\theta, I_{n_x})}{\partial \text{vec}T})|_{\theta=\theta_0}$

### 5.0.1. *Characterization of the Identified Set using Macroeconomic Data.*

As already shown, the identified set is defined by the conditional moment inequalities in 24,

$$\Theta_I \equiv \left\{ \theta \in \Theta : \mathbb{E}_t(Y_t^o - C(\theta)\hat{X}_{t|t-1} | \sigma(\mathcal{F}_{t-1})) \geq 0 \right\}$$

In all of the identification results we assume the existence of appropriate random variables to construct unconditional moment restrictions from the conditional moment inequalities. Such instruments can be either past data or past state variables constructed with the Kalman filter. By construction, the latter are uncorrelated with current information, but they might be noisy. In the case of additional moment conditions, a.k.a *supernumerary*<sup>15</sup>, we need to show conditions under which they can further reduce the identified set, as  $\Theta'_I$  will now satisfy multiple moment inequalities. Although the work of [Bontemps, Magnac, and Maurin \(2012\)](#) deals with linear models, we cannot use their results here for several reasons. First, the characterization of the identified set is done through the support function of  $\Theta_I$ <sup>16</sup>, which requires the identified set to be convex<sup>17</sup>. In our case, since the stability conditions and the cross equation restrictions introduce on  $\mathcal{L}(\theta)$  and therefore nonlinear restrictions on  $\Theta$ , the identified set is not necessarily convex and its geometry is difficult to be known a priori. Second, we deal with moment conditions of general form and therefore additional conditions arising from more instruments in an IV setting is just a special case. We nevertheless also provide an adapted *Sargan* condition. Third, point identification cannot in general occur despite the moment inequalities as we deal with typically continuously distributed random variables.

<sup>15</sup>We borrow this term from [Bontemps, Magnac, and Maurin \(2012\)](#).

<sup>16</sup>The support function, which is  $\sup(q^T \Theta)$ ,  $\forall q \in \mathbb{R}^{n_\theta}$ , can fully characterize any convex set  $\Theta$ .

<sup>17</sup>One can alternatively work with the Aumann expectation of the non-convex random set, which is always convex. Nevertheless, this requires minimizing the support function of the random set for any value of  $\theta$ , which is costly in high dimensional settings like ours ([Bontemps and Magnac, 2017](#)).

Recall that the number of moment conditions we use for estimation depend on the number of observables. Assumption **LCI-6** requires that there has to be enough (or the right kind) of observables such that a rank condition is satisfied. In our case, the number of observables used determines the number of first order conditions used for estimation. The minimum number ( $r$ ) of observables required such that conditional identification is achieved ( Lemma 4 is satisfied) maps to the necessary first order conditions. For example, if we have  $Y_1$  and  $Y_2$  to estimate the model, and we only need  $Y_1$  to conditionally identify  $\theta$ , then the  $n_\theta \times 1$  first order conditions arising from  $Y_1$  will be the necessary conditions. The rest of the conditions, i.e. those arising from  $Y_2$  are then supernumerary.

Notice that given the structure of the class of models we consider in this paper, it is straightforward to find a re-parameterization that restores moment equalities, that is, for every observable  $Y_t^o$ ,  $\exists U_t : \mathbb{E}U_t\phi(\mathbf{Y}_{t-1}) \equiv U \in \mathbb{R}^+$  that satisfies the restriction  $\mathbb{E}(Y_t^o - C(\theta)^*\hat{X}_{t|t-1})\phi(\mathbf{Y}_{t-1}) - U = 0$ . Then, for  $\mathbf{U} \equiv (U_1, U_2..U_r)'$ , the following set determines the map from  $U_I$  to  $\Theta_I$ :

$$(25)(\Theta_I, \mathbf{U}_I) := \left\{ (\theta, u) \in (\Theta, \times_r \mathbb{R}^+) : \mathbb{E}(Y_t^o - C(\theta)^*\hat{X}_{t|t-1})\phi(\mathbf{Y}_{t-1}) - U = 0 \right\}$$

**Proposition 4.**  $\Theta_I$  *when*  $n_y = r$

*Given correct specification of  $M_f$ , Lemma 4 and optimal instrument  $\phi^*(\mathbf{Y}_{t-1})$ , definition 25 implies the following:*

- $\exists!$  invertible mapping  $\mathcal{G} : \Theta_I = \mathcal{G}^{-1}(U) \cap \Theta$
- $\Theta_I$  is sharp

*Proof.* See Appendix

□

Note that the proposition applies to the infeasible case of using optimal instruments. If the case of non-optimal but valid instruments, the set is again as sharp as possible, up to the information loss implied by the non-optimality of the instrument<sup>18</sup>. It is also important to note that our characterization is based on the fact that the unobserved endogenous variables are integrated out using the Kalman filter. This implies that the only unobservables we need to deal with in the characterization of  $\Theta_I$  is the aggregate shocks in the economy. Moreover, the set is robust to individual heterogeneity as long as it vanishes on aggregate. More importantly, given correct specification,  $U$  is non zero if and only if there is a positive mass of agents that indeed face frictions.

Having established the determination of  $\Theta_I$  using the minimum number of observables, we turn to the case of using a larger number of observables, including non-macroeconomic variables. Let  $m_{\alpha,t}(\theta)$  and  $m_{\beta,t}(\theta)$  denote the necessary and super-numerary moment functions for identifying  $\Theta$ , where  $\mathbb{E}(m_{\alpha,t}(\theta)|\mathbf{Y}_{t-1}) = \mathcal{V}_\alpha(\mathbf{Y}_{t-1}) \in [\underline{\mathcal{V}}_\alpha(\mathbf{Y}_{t-1}), \bar{\mathcal{V}}_\alpha(\mathbf{Y}_{t-1})]$  and  $\mathbb{E}(m_{\beta,t}(\theta)|\mathbf{Y}_{t-1}) = \mathcal{V}_\beta(\mathbf{Y}_{t-1}) \in [\underline{\mathcal{V}}_\beta(\mathbf{Y}_{t-1}), \bar{\mathcal{V}}_\beta(\mathbf{Y}_{t-1})]$ . For notational brevity, we drop the dependence on  $\mathbf{Y}_{t-1}$  hereafter.

Comparing these general bounds to the ones implied by the model equilibrium restrictions,  $\underline{\mathcal{V}}_\alpha = \underline{\mathcal{V}}_\beta = 0$  and  $\bar{\mathcal{V}}_\alpha = \bar{\mathcal{V}}_\beta = \infty$ . Due to the boundedness of  $\Theta$  and the cross-equation and stability restrictions, the effective lower and upper bounds are likely to lie strictly within  $[0, \infty)$  for every moment condition. Let  $\phi(\cdot)$  be any  $\mathbf{Y}_{t-1}$ -measurable function for which  $\hat{m}_{\alpha,t}(\theta) := m_{\alpha,t}(\theta)\phi_t$  and  $\hat{m}_{\beta,t}(\theta) := m_{\beta,t}(\theta)\phi_t$ ,  $\hat{\mathbf{m}}_\alpha(\theta)$  and  $\hat{\mathbf{m}}_\beta(\theta)$  the corresponding vectors, and  $\bar{\mathbf{m}}_\alpha(\theta)$  and  $\bar{\mathbf{m}}_\beta(\theta)$  the vector means.

Let  $W$  be a real valued, possibly random weighting matrix, diagonal in  $(W_\alpha, W_\beta)$ . Furthermore, denote by  $Q_\alpha$  the conditional expectation operator when conditioning on  $W_\alpha^{\frac{1}{2}T} \hat{\mathbf{m}}_\alpha(\theta)$  and by  $Q_\alpha^\perp$  the residual.

---

<sup>18</sup>Recent work in the literature proposes constructing instrument functions to avoid this information loss, see for example [Andrews and Shi \(2013\)](#). We nevertheless do not pursue this in this paper.

The following proposition specifies the additional restrictions that need to be satisfied by the additional moment conditions such that the size of the identified set becomes smaller. This is a general result, and it applies both to conventional cases where  $n_Y > r$  or when data on non-macroeconomic variables is used.

**Proposition 5. *The Identified Set with Multiple Conditions Given***

$$(1) \Theta_I \neq \emptyset$$

$$(2) U_t \in [\underline{U}, \bar{U}], \text{ with } \underline{U} \equiv (W_\alpha^{\frac{1}{2}T} \underline{\nu}_\alpha + W_\beta^{\frac{1}{2}T} \underline{\nu}_\beta)\phi_t, \bar{U} \equiv (W_\alpha^{\frac{1}{2}T} \bar{\nu}_{\alpha t} + W_\beta^{\frac{1}{2}T} \bar{\nu}_\beta)\phi_t$$

Then  $\Theta'_I \subset \Theta_I$  iff

$$(26) \quad \mathbb{E}Q_\alpha^\perp \left( W_\beta^{\frac{1}{2}T} \hat{m}_{\beta,t}(\theta) - \hat{U}_t \right) = 0$$

*Proof.* See the Appendix □

The main argument behind Proposition 5 is the following. Suppose that the necessary moment conditions have no common information with the supernumerary conditions and that  $W = I_{n_\alpha+n_\beta}$ . From the minimization of  $\mathbb{E}\frac{1}{2}(\bar{\mathbf{m}} - \bar{U}_t)^T(\bar{\mathbf{m}}(\theta) - \bar{U}_t)$  where  $\bar{\mathbf{m}}(\theta) \equiv (\bar{\mathbf{m}}_\alpha(\theta), \bar{\mathbf{m}}_\beta(\theta))^T$ , the first order condition is

$$\mathbb{E}(\bar{\mathbf{m}}_\alpha(\theta) + \bar{\mathbf{m}}_\beta(\theta) - \bar{U}_t) = 0$$

which for  $\bar{U}_\alpha \equiv \overline{Q_\alpha^\perp U_t}$  can be rewritten as  $\mathbb{E}((\bar{\mathbf{m}}_\alpha(\theta) - \bar{U}_\alpha) + (\bar{\mathbf{m}}_\beta(\theta) - \bar{U}_\alpha^\perp)) = 0$ . By construction the two parts of the left hand side of the expression are independent, and therefore both have to be zero.

$$(27) \quad \mathbb{E}(\bar{\mathbf{m}}_\alpha(\theta) - \bar{U}_\alpha) = 0$$

$$(28) \quad \mathbb{E}(\bar{\mathbf{m}}_\beta(\theta) - \bar{U}_\alpha^\perp) = 0$$

Notice that, by construction, the set of necessary moment conditions in 5.3. must have full rank, and this establishes a one-to-one mapping from  $\Theta$  to the domain of variation of  $U_t$ ,  $[\underline{U}(\mathbf{Y}_{t-1}), \bar{U}(\mathbf{Y}_{t-1})]$ . Thus, there exists an inverse operator  $\mathcal{G}_\alpha$  such that  $\theta = \mathcal{G}_\alpha(U_t, \mathbb{P})$ . Plugging this expression for  $\theta$  in 5.4, we get that

$$\mathbb{E}\bar{m}_\beta(\mathcal{G}_\alpha(U_t, \mathbb{P})) = \mathbb{E}\bar{U}_\alpha^\perp$$

This is a restriction on the values that  $U_t$  can take in addition to the ones implied by the necessary conditions. A restriction on  $U_t$  implies a restriction on the admissible  $\Theta_I$  given the one-to-one relationship in 5.3. Notice that when the supernumerary conditions do not add any additional information, i.e.  $m_{\alpha,t}(\theta) \equiv m_{\beta,t}(\theta)$ , the restriction collapses to  $Q_\alpha = Q_\alpha^\perp = \frac{1}{2}$ . Given these identification results, we can analyze identification arising from any inequality restriction, and therefore any additional type of information can be potentially analyzed. Below, we discuss and formally show how qualitative survey data can provide additional restrictions that are informative about aggregative models of economies with frictions. Such information constrains further the stochastic properties of  $\lambda_t$ , and therefore the size of  $\Theta_I$ .

## 6. THE USE OF ADDITIONAL INFORMATION

Additional data can be potentially linked to dynamic equilibrium models by augmenting the observation equation with moment restrictions. The latter can be motivated by the fact that not all data are explicitly modeled using structural equations, but they can nevertheless be used to sharpen economic and econometric inference, especially for unobserved processes. An example of such a process is the distortion  $\lambda_t$ . We focus on qualitative survey data because, as we illustrate, they contain distributional information that can be linked to the aggregate model and are informative about  $\lambda_t$ . This is especially true for economies with frictions, as the proportion of agents whose

behavior is distorted is an important statistic. An example of a distributional statistic that has been routinely used to judge the validity of *complete* models is the proportion of firms which cannot change prices in New Keynesian models featuring a Calvo adjustment mechanism. Our focus is on more general information that is contained in qualitative surveys, and given that we deal with *incomplete* models, this information is able to discard a set of economic mechanisms, and not just one. We thus view our method as a significant generalization of the treatment of microeconomic information in dynamic equilibrium models.

Qualitative survey data are usually available in the form of aggregate statistics, where aggregation is performed over categories of answers to particular questions. As in micro-econometric studies, the categorical variable is a function of a continuous latent variable. In a structural context, there is a measurable mapping from the categories of answers to the random variables relevant to the decision of each agent. For example, if the question is of the type "How do you expect your financial situation to change over the next quarter" and the answer is trichotomous, i.e. "Better", "Same" and "Worse", then the answers map to a set of partitions of the end of period assets  $a_{t+2}$ :  $a_{t+2} \in [a_{t+1} - \epsilon, \infty)$ ,  $a_{t+2} \in (a_{t+1} - \epsilon, a_{t+1} + \epsilon)$  and  $a_{t+2} \in (-\infty, a_{t+1} - \epsilon]$ . For this interpretation to hold, we need to assume that agents report their states or beliefs truthfully. Furthermore, denote by  $\{S_{i,k,t}\}_{i \leq N}$  the survey sample over a period of length  $T$ , for the  $i$ th respondent. Let  $C_l^k$  be the  $l^{\text{th}}$  categorical answer to question  $k$  and  $\hat{\xi}_{i,t,k}$  the respondent's choice. Given some weights on each category  $w_l$ , the available statistics are of the form:  $\hat{B}_t^k = \sum_{l \leq L} w_l \sum_{i \leq N} 1(\hat{\xi}_{i,t} \in C_l^k)$ . Given truth telling, we can map the answer to agent beliefs i.e. there exists a mapping  $h : \hat{\xi}_{i,t} \rightarrow \xi_{i,t} \equiv \mathcal{E}_{i,t}(x_{i,t+1} | x_{i,t}, X_t)$ . Moreover, since the conditional expectation is a function of the information set, let  $B_l \equiv \{(x_{i,t}, X_t) \in \mathbb{R}^{2n_x} : h(x_{i,t}, X_t) \in C_l^k\}$ , that is,  $B_l$  belongs to the partition of the support of individual  $x_{i,t}$  (or aggregate  $X_t$ )

that corresponds to category  $C_l^k$ . Consequently, the survey statistic has the following theoretical form, which can be linked to the model:

$$(29) \quad \hat{\mathcal{B}}_t^k = \sum_{l \leq L} w_l \sum_{i \leq N} w_i 1(\mathcal{E}_{i,t}(x_{t+1,i}|x_{i,t}, X_t) \in B_l)$$

For every  $Y_t^o$  with a corresponding model based conditional expectation  $Y_t^m \equiv \mathbb{E}(Y_t^o|M, \mathcal{F}_t)$ , Proposition 6 illustrates the additional restrictions implied by  $\hat{\mathcal{B}}_t^k$ . For these to hold, we need to make the following assumptions:

. **ASSUMPTION -S**

- (1) Let  $\hat{\mathcal{B}}_t^k$  defined as in (29) and an aggregate shock vector  $Z_t$  of length  $p > 0$ , with innovation  $\epsilon_t|\mathcal{F}_{t-1} \sim (0, \Sigma_e)$ .
- (2)  $\lambda_{i,t}$  is conditionally linear.
- (3) Truthful response such that  $\hat{\mathcal{B}}_t^k \rightarrow B_t \in (0, 1]$ .
- (4)  $\hat{R}_t^m := \left( \frac{\mathbb{E}(x_{i,t-1}^m|x_{i,t} \in \mathcal{X}_t^{*,C})}{\mathbb{E}(x_{i,t-1}^m|x_{i,t} \in \mathcal{X}_t^*)} - 1 \right) \mathbb{P}(x_{i,t} \in \mathcal{X}_t^{*,C})$  is monotone in  $\mathcal{X}_t^{*,C}$  for all  $m = 1..n_x$ .

The assumption of conditional linearity for  $\lambda_{i,t}$  enables us to claim that once the distortion is activated, it is a linear function of the states and the shocks, e.g.  $\lambda_{i,t} := \lambda_1 x_{i,t-1} + \lambda_2 z_{i,t}$ . We motivate assumption S-4 after presenting the following proposition:

**Proposition 6.** *Under Assumption-S, the following restriction holds,  $\mathbb{P}_{X_{t-1}} - a.s.$*

$$\mathbb{E}(Y_t^o|X_{t-1}) \leq P(\theta_1, 0)X_{t-1} + \lambda^1 X_{t-1} \odot \mathbb{E}(B_t|X_{t-1})$$

where  $P(\theta_1, 0) \equiv C(\theta)$ .

*Proof.* See Appendix A

□

Notice that in the proof, we define  $(\mathcal{X}_t^*, \mathcal{Z}_t^*)$  as the time dependent subset of the support such that  $\lambda_{i,t} \neq 0$ . In models of heterogeneous agents these boundaries usually depend on aggregate states and shocks. Therefore, this inequality will in general be strict unless  $B_t = 0$ .

Moreover, in the proof we establish two facts. First, since the equilibrium conditions of the model with frictions depend on subjective conditional expectations, and the model with frictions is a smooth perturbation of the frictionless model, any probability statement on the subjective expectations translates to a probability statement on  $\mu_{i,t}$ . Second, any probability statement on  $\mu_{i,t}$  is a probability statement on the solution of the model, and therefore on  $\lambda_{i,t}$ . Given representation (29), qualitative survey data have information on the quantiles of subjective conditional expectations of the agents. Therefore, survey data relate directly to the conditional probability of observing a friction,  $\mathbb{P}_t(\lambda_{i,t} \geq 0)$ , which implies the above restrictions. The latter cannot in general deliver point identification, as the vector  $\hat{\mathbf{R}}_t = (\hat{R}_t^1, \hat{R}_t^2 \dots \hat{R}_t^{n_x})$  is hard to pin down unless specific assumptions are made. Monotonicity, which is a mild assumption, provides a bound.

Given Proposition 6, we can also establish the following corollary result regarding identification using survey data.

**Corollary 7.** *Identification with Survey Data*

- (1) If  $\hat{\mathcal{B}}_t^k \rightarrow B_t \in (0, 1)$ ,  $\Theta'_I \subset \Theta_I$ .
- (2) *Impossibility of point identification:* When  $B_t \neq 0$ ,  $\Theta_I$  is not a singleton

*Proof.* See Appendix □

Corollary 7 has several implications. First, the set of admissible structures becomes smaller, and we can therefore make more precise statements regarding parameters and

conditional predictions. Second, using identified set and the definition of  $\lambda(Y_t, \theta)$ , a plug-in set estimate of the average distortion (wedge) in a macroeconomic variable  $Y_t$  is  $\mathbb{E}\lambda(Y_t, \theta_I(Y))$ . These estimates of distortions can be used in several ways, which we explore in the next section.

## 7. INFERENCE BASED ON $\lambda(\Theta_I)$

We first discuss briefly how  $\Theta_I$  and  $\lambda(\Theta_I)$  can be obtained in practice. Given the quasi-structural framework, performing inference for  $\Theta_I$  rather than  $\theta_0$  seems a natural thing to do. Many of the parameters in dynamic macroeconomic models are semi-structural, and therefore  $\theta_0$  itself does not have a very specific economic interpretation. We therefore use methods appropriate for constructing consistent estimators for the identified set (IdS) and confidence sets (CS) for the IdS. In the frequentist literature, several methods have been proposed for a general criterion function like subsampling [Chernozhukov, Hong, and Tamer \(2007\)](#); [Romano and Shaikh \(2010\)](#) or the bootstrap for moment inequality models [Bugni \(2010\)](#). Nevertheless, in macroeconomic models parameter dimensions are high and pointwise testing is a vastly inefficient way to construct confidence sets, so we focus on computationally tractable methods to perform statistical inference using Markov Chain Monte Carlo algorithms (MCMC).

When  $\theta_0$  is point identified and root- $n$  estimable, [Chernozhukov and Hong \(2003\)](#) have proposed to use simulation methods to do inference on  $\theta$  using the quasi-posterior distribution, which is constructed using the relevant loss function  $L_n(\theta)$  that defines  $\theta_0$ . Generalized Method of Moment (GMM) class of estimators can be easily embedded in this case. Denoting by  $\pi(\theta)$  the prior distribution, simulation draws  $\{\theta^j\}_{j \leq L}$  are obtained from

$$\Pi_n(\theta|\mathbf{Y}) \equiv \frac{\exp(TL_n(\theta))}{\int_{\Theta} \exp(TL_n(\theta))d\pi(\theta)}$$

and upper and lower  $100(1 - \alpha)/2$  quantiles are used to conduct inference that has valid frequentist properties. Nevertheless, once point identification fails, one has to consider adjusting the method to accommodate for this.

More particularly, let  $L_n(\theta) = n^{\frac{1}{2}}q_n(\theta)'_+W_n n^{\frac{1}{2}}q_n(\theta)_+$  be the criterion function to be minimized, where  $q_n(\theta)$  are the moment functions to be used and  $q_n(\theta)_+ \equiv \max(q_n(\theta), 0) = \min(-q_n(\theta), 0)$ . If  $L_n(\theta)$  is stochastically equicontinuous, that is, there exists a  $\Delta_n(\theta_0)$  and  $J_n(\theta_0)$  such that  $L_n(\theta)$  admits a quadratic expansion <sup>19</sup> then this allows us to assume a Central Limit Theorem (CLT) on  $\Delta_n(\theta_0)$ . Redefining  $Q_n(\theta) \equiv q_n(\theta) - b$  where  $b$  is the bias term and  $\tilde{L}_n(\theta, b) = n^{\frac{1}{2}}Q_n(\theta)'W_n n^{\frac{1}{2}}Q_n(\theta)$ , we assume that the following CLT holds

$$(30) \quad V(\theta_0, b_0)^{-\frac{1}{2}}\tilde{\Delta}_n(\theta_0, b_0) \xrightarrow{d} N(0, I)$$

Estimating equations arising in DSGE models involve smooth functions. This is particularly true if we focus on determinate equilibria, so that model specification is uniform across the parameter space. Combining them with smooth instrument functions is enough to guarantee stochastic equicontinuity.

Given <sup>30</sup>, [Liao and Jiang \(2010\)](#) show that when using a cutoff rate  $\nu_n$  such that  $1 < \nu_n < n$ , and defining  $\mathcal{A}_n := \{\max_{\vartheta} \ln(\Pi_n(\vartheta|\mathbf{Y})) - \ln(\Pi_n(\theta|\mathbf{Y})) \leq \nu_n\}$  then  $d_H(\mathcal{A}_n, \Theta_I) \rightarrow_p 0$  where  $d_H(\cdot, \cdot)$  is the Hausdorff distance between two sets <sup>20</sup>. Moreover, the rate of convergence of the pseudo-posterior density outside the identified region is exponential. As is evident, in finite samples the identified will depend on  $\nu_n$ , so some robustness checks are required. In addition, since our criterion function vanishes when

---

19

$$L_n(\theta) = L_n(\theta_0) + (\theta - \theta_0)' \Delta_n(\theta_0) - \frac{1}{2}(\theta - \theta_0)' n J_n(\theta_0) (\theta - \theta_0) + R_n(\theta)$$

<sup>20</sup> $d_H(\cdot, \cdot) = \max[\sup_{a \in A} \inf_{b \in B} \|b - a\|, \sup_{b \in B} \inf_{a \in A} \|a - b\|]$

$\theta \in \Theta_I$ , we do not need to adjust  $\nu_n$  as suggested in [Chernozhukov, Hong, and Tamer \(2007\)](#).

With regards to constructing confidence sets for the identified set, which is what we do in the empirical application, recent work by [Chen, Christensen, O’Hara, and Tamer \(2016\)](#) provides a computationally attractive procedure to construct  $\Theta_\alpha^{CS}$  for the IdS and functions of it that have correct coverage from a frequentist perspective. Compared to [Chernozhukov and Hong \(2003\)](#); [Liao and Jiang \(2010\)](#), cutoff values are based on quantiles of draws of the loss function  $L_n(\theta)$  rather than draws from the quasi-posterior of the parameter vector and there is no maximization involved. Also, results extend to models with singularities, that is models in which a local quadratic approximation involves a non vanishing singular component, and parameters are not root- $n$  estimable. This is useful in case one wants to derive bounds based on extensions to non-parametric treatments of the survey based bounds. Although we find the method suggested very appealing, we apply it only in the case of estimating frictions. Proving validity for the testing procedure we will propose in this section is not trivial as it involves combining two independent MCMC chains, and we leave this interesting research for the immediate future.

We next motivate the proposed test that utilizes  $\lambda_Y(\Theta_I)$ . When a *complete model* is estimated, the pseudo true vector  $\theta_0$  is point identified. Given its value, we can estimate the predicted level of friction in each of the macroeconomic variable,  $\mathbb{E}\lambda_Y(\theta_0)$ . Misspecification of this model may imply that the predicted friction does not lie in the identified set of distortions. Therefore, the distance from  $\mathbb{E}\lambda_Y(\Theta_I)$  becomes a sufficient statistic to judge whether the suggested model is properly specified. We propose a Wald statistic that tests whether the expected distance from the point estimate for the wedge from the parametric model to the identified set of wedges is different than zero for all (or some of) the observables. Since survey data provide

more information on the wedge, this increases power to reject non-local alternatives. In addition, survey data regularizes the test, as in the absence of additional data the distribution degenerates and requires non-standard inference.

7.0.1. **A Test for Parametric Models of Frictions.** Letting  $H_0 : \theta_p \in \Theta_s$  and  $H_1 : \theta_p \notin \Theta_s$ , the proposed statistic is as follows:

$$\mathcal{W}_t = \left( \sqrt{T} \inf_{\lambda_s \in \hat{\Theta}_s} \|\mathcal{V}^{-\frac{1}{2}}(\lambda_s - \lambda_p^*)\| \right)^2$$

where  $\lambda$  is the estimated friction obtained using either the identified point in the parametric model case,  $\lambda_p$ , or the identified set in the robust case,  $\lambda_s$ . Individual frictions are weighted by their respective estimate of standard deviation. The statistic measures the Euclidean distance between the wedge that arises in the parametric model, and the set of admissible wedges, adjusted for estimation uncertainty. Under the following conditions, the test is consistent and has asymptotic power equal to one against fixed alternatives.

. **ASSUMPTION -R** (Regularity conditions) Let  $D$  and  $q$  be  $Y$ -measurable functions, continuous in  $\theta$  w.p.1 and  $\tilde{q}(\cdot; \theta) \equiv \tilde{q}(\cdot; \theta) - \bar{q}(\cdot; \theta)$  such that:

- (1) For any  $\mathbf{Y}$ , and any  $\theta \in \Theta$ ,  $\sqrt{T}\tilde{q}(\mathbf{Y}; \theta) \rightarrow_d \mathcal{N}(0, \Omega)$
- (2)  $\sup_{\theta \in \Theta} D_n(\theta) \rightarrow_p D$  where  $D$  is positive definite

**Proposition 8.** *Consistency and Power.* Under assumption **R**:

- (1)  $T\mathcal{W}(\theta_p, \Theta_s) \xrightarrow{d} \|\sum_{j=1..p} \omega_j \mathcal{N}(0, I_p)\|^2$
- (2) Given critical value  $c_\alpha$  with  $\alpha \in (0, 1)$ ,
  - (a) Under  $H_0$ :  $\lim_{T \rightarrow \infty} \mathbb{P}(T\mathcal{W}(\theta_p, \Theta_s) \leq c_\alpha) = 1 - \alpha$
  - (b) Under  $H_1$ :  $\lim_{T \rightarrow \infty} \mathbb{P}(T\mathcal{W}(\theta_p, \Theta_s) \leq c_\alpha) = 0$

*Proof.* See Appendix

□

In the supplemental material we show that using the non-parametric block bootstrap to compute  $c_\alpha$  is valid, and we illustrate through a measurement error example that the bootstrap distribution coincides with the asymptotic distribution. Thorough examination of the performance of this test in small samples is a very interesting avenue of research, which deserves separate treatment. Another equally interesting topic for future research is to investigate which criterion function would deliver robustness of the testing procedure to possible misspecification of the benchmark model, which we have assumed away. See [Ponomareva and Tamer \(2011\)](#) for the issue of misspecification in moment inequality models.

Finally, we briefly discuss the *refutability* of the candidate model, in the sense of [Breusch \(1986\)](#). The Null hypothesis, tests whether the point identifying restrictions on  $\lambda_Y$  implied by the candidate model are satisfied on  $\lambda_{Y,s}$ , where the latter is only set identified. Since  $(\lambda_{Y,s}, \lambda_{Y,p})$  is constructed using  $\theta_1$  as identified using the frictionless and complete model respectively (See Proposition 1 for the definition of  $\theta_1$ ),  $\lambda_{Y,p}$  is a function of the additional parameters indexing the CM,  $\theta_2$ . In the absence of cross parameter (equation) restrictions on  $\Theta_1 \times \Theta_2$ ,  $\lambda_{Y,p}$  would necessarily lie within  $\lambda_{Y,s}$  as the latter would be consistent with any  $\theta_2 \in \Theta_2$ . However, in the presence of cross equation restrictions,  $(\theta_1, \theta_2)$  lie in a strict subset of  $\Theta_1 \times \Theta_2$ . Thus, the completion  $\lambda_{Y,p}$  may not necessarily lie within the family of completions,  $\lambda_{Y,s}$ . Equivalently,  $\exists \lambda \in \lambda_s$  that is not observationally equivalent to  $\lambda_p$ . In this sense, the candidate model is *refutable* only when it imposes restrictions on  $\lambda_y := U_y(\theta_1, \theta_2)$  for any observable  $Y$ , which implies restrictions on the reduced form. This is indeed true in the context of incomplete general equilibrium models; taking the unconditional expectation in condition (23) of Proposition 1, implies that  $(\Theta_1, \Theta_2)$  is not variation free and therefore  $U_y(\theta_1, \theta_2)$  is restricted.

## 8. APPLICATION: ESTIMATING FRICTIONS IN THE SPANISH ECONOMY

We apply the methodology to the case of Spain, where financial frictions have arguably played a significant role during the last decade. The benchmark economy we use features *some* frictions and since it is standard, we will directly introduce the log-linearized conditions. We consider a small open economy with capital accumulation, along the lines of [Smets and Wouters \(2007\)](#) and [Gali and Monacelli \(2005\)](#). There are households, intermediate good firms, final good firms, government expenditure, and a foreign sector which is composed by infinitesimal symmetric economies.

The type of frictions we allow in the baseline model are those we do not have sufficiently informative survey data to implement our methodology. Thus, we keep the parametric Calvo type of friction in the wage setting by labor unions and in the price setting behavior of firms. All other frictions are going to be semi-parametrically characterized. We thus remove capital adjustment costs, and therefore Tobin's  $q$  becomes constant. This implies that the arbitrage condition between capital and bonds has no dynamics in the benchmark model.

Let  $X_1^o$  denote the vector of variables that enter the moment equalities,  $X_2^o$  the vector of variables used in the moment inequalities and  $Z$  the vector of instruments. Model predictions are denoted with superscript ' $m$ '. The conditions we use are therefore:

$$\mathbb{E}((X_{1,t}^o - X_{1,t}^m) \otimes Z_t) = 0$$

$$\mathbb{E}((X_{2,t}^o - X_{2,t}^m) \otimes Z_t) \leq 0$$

The variables employed in estimation are Non Government Consumption expenditure (C), Hours (H), Inflation ( $\pi$ ), Investment (I), Gross Domestic Product (Y), Wages (W), and the EONIA rate (R). Real variables are in per capita terms. We use lagged

values of output, consumption and investment as instruments, as all positively depend on the capital stock which is unobserved. We estimate the model using business survey data from Spain, collected by the European Commission. Our sample period covers 1999 $Q_1$  to 2013 $Q_4$ . Detailed information on this survey data can be found at: [http://ec.europa.eu/economy\\_finance/db\\_indicators/surveys/index\\_en.htm](http://ec.europa.eu/economy_finance/db_indicators/surveys/index_en.htm). We include responses to questions 8F4 and 8F6, relating to capital adjustment costs and financial constraints to production capacity. To isolate the effects of using just the firm survey data, we do not use the consumer survey.

For the case of capital adjustment costs, we have restrictions on  $Y$  and  $I$  similar to those of the general example in Section 3. Financial constraints to productive capacity imply similar restrictions and lead to lower aggregate investment and output. Although we do not use the consumer survey, we also allow for consumer credit constraints. The later imply negative distortions to household consumption, and given that hours worked are complementary to consumption, they also imply negative distortions to labor supply and output.

We therefore choose  $X_{1,t}^o \equiv (W, \pi, R)$  and  $X_{2,t}^o \equiv (Y, I, H, C)$ . We estimate the model using the 2-block RW-MCMC with uniform priors for all parameters<sup>21</sup>. Since the survey based moment restrictions require estimating the nuisance parameters  $\lambda^1$  (see Proposition 6), for a given draw of  $\Theta_1$ , we obtain  $\hat{\lambda}^1$  by ordinary least squares. This approximation can, in theory, weaken the informativeness of the additional restrictions but we do not observe this in our results.

**8.1. Details of the exercise.** We obtain the 95% confidence sets of wedges, that is, a range for the standardized wedge in each observable that is consistent with both macroeconomic and survey data,  $\lambda_y \equiv (\mathbb{E}V_{y|\hat{x}}\mathbb{E}V_{\hat{x}})^{-1}\mathbb{E}\hat{X}_{t,t-1}(Y_t - C(\Theta_I))$ , which

---

<sup>21</sup>We keep the last 200000-300000 draws for inference.

under linearity is equivalent to  $(\mathbb{E}V_{y|\hat{x}})^{-1}\lambda_y$ <sup>22</sup>. These estimates are the empirically relevant wedges that any model featuring financial frictions and adjustment costs should produce over the whole sample. We plot them (in red) in the left panel of Figures 1-4, in Appendix A.

We employ a complete model (**CM**) featuring [Bernanke, Gertler, and Gilchrist \(1999\)](#) (BGG) type of frictions (similar to [Graeve \(2008\)](#)) and estimate it using the same macroeconomic aggregates as the IM using full information MCMC, where we also rely on [Chen, Christensen, O'Hara, and Tamer \(2016\)](#) to obtain  $\Theta_{\alpha,CM}^{CS}$ .

Despite that the model features financial frictions on the firm side (the "entrepreneurial sector"), and not borrowing constrained consumers, its implications can still be tested against the incomplete model (**IM**), which is robust to both types of frictions.

We plot (in blue) the corresponding confidence sets for the wedges that are consistent with the complete model. To facilitate the comparison, we also produce an alternative comparison based on QQ plots, which gives information on the support of each of these estimates. Note that in some of the cases the support is completely different.

In Table 1 (Appendix A), we display the corresponding confidence sets for  $\Theta$  using survey data. Confidence sets for  $\Theta_{IM}$  without using survey data and the corresponding wedges can be found in Appendix B.

A key observation is the fact that the sign of the wedge confidence sets that correspond to the restrictions imposed to identify the IM are in accordance with the sign of the distortions identified with the CM. While both the CM and IM identify negative distortions to output, the former identifies a significantly lower level of distortions. As is typical, empirical versions of accelerator type of models ignore the output costs

---

<sup>22</sup>We actually compute  $\lambda_t$  using quadratic programming as in the second section of the online Appendix (page 70 of this script) and then take the average. The result is equivalent to a standardized wedge.

(bankruptcy costs) as they are deemed to be numerically unimportant. Some of the neglected variation is indeed captured by the high estimates of  $\phi_{p,CM}$  relative to  $\phi_{p,IM}$ .

Regarding inflation, using the CM we cannot reject zero frictions to  $\pi_t$ , while the IM identifies significantly negative distortions. Moreover, the CM identifies a much higher slope of the Phillips curve<sup>23</sup>, which is consistent with its "inability" to generate negative distortions to inflation.

With regard to the nominal interest rate, both the CM and the IM identify negative distortions (with no evidence of different magnitudes). Mechanically, this is due to the Taylor rule that pins down  $r_t$ . Nevertheless, there is also a deeper insight which comes from steady state reasoning. In an economy with uninsured idiosyncratic risk on both firms and consumers, the steady state interest rate is lower than in the case of complete markets, and as shown by [Angeletos \(2007\)](#), this does not contradict negative distortions to capital accumulation due to the presence of a risk premium.

Furthermore, we cannot reject the hypothesis that the CM is consistent with zero distortions to  $c_t$ , while the IM identifies negative distortions. Moreover, the estimates of risk aversion are much lower in the IM compared to the CM. One of the reasons is the fact that the CM ignores credit constraints to consumers, and therefore requires much higher estimates in risk aversion to match the corresponding negative distortions to consumption.

With regard to hours worked, both the IM and the CM identify negative distortions which are statistically different. Since  $\sigma_{c,IM}$  is much smaller than  $\sigma_{c,CM}$ , negative distortions to consumption are consistent with less distortions to hours in the IM. Moreover, the CM identifies a slightly higher (inverse) Frisch elasticity ( $\sigma_{l,CM}$ ) and

<sup>23</sup> $\kappa := \frac{(1-\iota_p\xi_p)(1-\xi_p)}{(1+\beta\iota_p)(\xi_p(1+(\phi_c-1)\epsilon_p))}$  where  $\epsilon_p$  (Kimball aggregator) is calibrated to 10.

does not reject zero distortions to  $w_t$ , while the IM identifies positive distortions. The latter observation is not robust to excluding the survey data as shown in Table 2.

Finally, the IM and CM identify the same negative distortions to investment. Since no marginal  $\Theta_{cs}$  has an extremely narrow support, this result can be possibly explained by the mapping from parameters to investment that is nearly not injective. This is not a problem per se, as the MCMC method we have used to obtain confidence sets in both the IM and the CM is robust to lack of identification.

The non overlapping standardized output, inflation and labour wedges provide enough evidence to reject the hypothesis that the SW-BGG model implies similar distortions to those identified by the IM. This finding also corroborates existing evidence of the lack of uniform fit for such a model over the entire business cycle (i.e. [Del Negro, B. Hasegawa, and Schorfheide \(2016\)](#) for the case of the US).

## 9. CONCLUSION

In this paper we propose a new inferential methodology which is robust to misspecification of the mechanism generating frictions in dynamic stochastic economies. We characterize wedges in equilibrium decision rules in a way which is consistent with a variety of generating mechanisms. We use the implied restrictions to partially identify the parameters of the benchmark model, and to obtain a set of admissible economic relationships. Moreover, we formally characterize set identification given the number of observables. Regarding the latter, beyond macroeconomic data, we show how qualitative survey data can provide additional distributional information, which is crucial in economies with ex post heterogeneous and frictions. Survey data provide a sufficient statistic for this information that would otherwise have to be generated by a model. We show how to exploit this additional information to test parametrically

specified models of frictions. We apply our methodology to estimate the distortions in the Spanish economy due to financial frictions and adjustment costs using an small open economy version of [Smets and Wouters \(2007\)](#) model and qualitative survey data collected by the European Commission. We investigate the adequacy of [Bernanke, Gertler, and Gilchrist \(1999\)](#) type of financial frictions.

In general, our work shows that adopting a robust approach to inference and using the information present in surveys is a fruitful way of dealing with lack of knowledge about the exact mechanisms generating frictions.

Our work is limited to dynamic *linear* economies and has focused on deviations of the representative agent approximation from the underlying heterogeneous agent economy. Possible extensions could study deviations based on a benchmark heterogeneous model. [Buera and Moll \(2015\)](#) have recently shown that the distortions generated in an aggregate equilibrium condition can depend on the type of heterogeneity present in the economy. Our approach is robust to this criticism. The methodology does not rest on *observing* residuals from representative agent frictionless economies - we just impose moment inequality restrictions theoretically motivated by deviations from the frictionless economy. In addition, if heterogeneity has additional implications for the form of these restrictions, they can be taken on board. In addition, we impose weak moment inequalities. This is important when heterogeneous distortions in decision rules cancel out. Furthermore, we impose restrictions implied by  $\mu_t$  on all of the variables and thus take general equilibrium effects into account. Nevertheless, future work could focus on investigating the robustness of our methodology in environments where some heterogeneity is ignored when imposing the identifying restrictions.

Finally, we have assumed that aggregated survey data are not systematically biased and strategy proof. Analysing the consequence for inference would be another fruitful avenue of research.

10. APPENDIX A

*Proof. of Proposition 1.*

Recall the representation for the model with frictions, that is,

$$G(\theta_1, 0)X_t = F(\theta_1, 0)\mathbb{E}_t(X_{t+1}|X_t) + L(\theta_1, 0)Z_t + \tilde{\mu}_t$$

Plugging in the candidate distorted decision rule:  $X_t^* = X_t^{f, RE} + \tilde{\lambda}_t$  and using that  $F(\theta_1, 0)P^*(\theta_1, 0) + G^*(\theta_1, 0) = 0$  and  $(R(\theta_1, 0)^T \otimes F(\theta_1, 0) + I_z \otimes (F(\theta_1, 0)P^*(\theta_1, 0) + G^*(\theta_1, 0)))\text{vec}(Q(\theta_1, 0)) = -\text{vec}(L(\theta_1, 0))$  we have the following condition:

$$\mathbb{E}_t(F(\theta_1, 0)\lambda_{t+1} - G(\theta_1, 0)\lambda_t + \tilde{\mu}_t) = 0$$

Note that in this proposition we let the econometrician's model variables (observables  $Y_t$  and unobservables  $Z_t$  coincide with a proper subset of  $X_t$  and  $Z_t$ . That is, setting  $\theta_2 = 0$  essentially eliminates some of the elements of  $(X_t, Z_t)$ . Furthermore, in the proposition we state that  $\lambda_t = \lambda_{t-1}\Gamma + \nu_t$  for some real-valued  $\Gamma \neq \mathbf{0}$ . Substituting for  $\lambda_t$  we get the condition stated in Proposition 1, that is:

$$(31) \quad \mathbb{E}_t(F(\theta_1, 0)\Gamma - G(\theta_1, 0))\lambda_t + \mu_t = 0$$

To motivate the assumption on the random variable  $\lambda_t$  notice that, the condition above essentially links the conditional mean of  $\lambda_t$  with that of  $\mu_t$ . Since  $\mu_t$  is by construction a linear function of  $X_t$ , then  $\mu_t$  is correlated with subsets of the information set of the agent,  $\sigma(\mathcal{F}_t)$ . Then, there exists a projection operator  $\mathbf{P}$  such that for any element  $H_t \in \mathcal{F}_t$ ,  $\mu_t = \mathbf{P}H_t + \mathbf{P}_\perp H_t = \mathbf{P}H_t + v_t$ . Projecting  $\mu_{t+1}$  on  $\mu_t$ :

$$\begin{aligned} \mathbf{P}(\mu_{t+1}|\mu_t) &= \mathbf{P}(\mathbf{P}H_{t+1} + v_{t+1}|\mathbf{P}H_t + v_t) \\ &= \mathbf{P}(\mathbf{P}H_{t+1} + v_{t+1}|\mathbf{P}H_t) + \mathbf{P}(\mathbf{P}H_{t+1} + v_{t+1}|v_t) \\ &= \mathbf{P}(H_{t+1}|H_t) \end{aligned}$$

This implies that, there exists a real valued matrix  $\tilde{\Gamma}$  such that  $\mu_{t+1} = \tilde{\Gamma}\mu_t + u_{t+1}$ . Substituting in the condition 9.1 and collecting all the errors in  $w_{t+1}$  (we can do this as they are not uniquely defined) we have that:

$$(F(\theta_1, 0)\Gamma - G(\theta_1, 0))\lambda_{t+1} = \tilde{\Gamma}(F(\theta_1, 0)\Gamma - G(\theta_1, 0))\lambda_t + w_{t+1}$$

Denoting by  $\tilde{C}$  the generalized inverse of  $C := F(\theta_1, 0)\Gamma - G(\theta_1, 0)$  we have that  $\lambda_{t+1} = \tilde{\Gamma}\lambda_t + C\tilde{C}w_{t+1}$ . Comparing with the proposed representation for  $\lambda_t$  we have that  $\Gamma = \tilde{\Gamma}$  and a non uniquely defined  $\nu_t$  that nevertheless satisfies  $\mathbb{E}(\nu_t) = 0$

**Note:** The assumption that the system can be casted in the expectational form to which we apply the method of undetermined coefficients is with no loss of generality. More elaborate methods like a canonical or Schur decomposition, can be used to obtain forward and backward solutions, and the existence of an incomplete rule would require arguments similar to the non zero determinant stability conditions for  $\Gamma(\theta)$  and  $\mu_t$ .  $\square$

*Proof. of Lemma 3*

Given  $\mathcal{V}_i(\mathbf{Y}_{t-1})$ , the condition in 5.1 can be rewritten as  $\mathbb{E}(\tilde{Y}_t^o - C(\theta)\hat{X}_{t|t-1})\phi(\mathbf{Y}_{t-1}) = 0$  where  $\tilde{Y}_t^o \equiv Y_t^o - \mathcal{V}(\mathbf{Y}_{t-1})$ . Given **LCI**, the Proposition 2-NS in [Komunjer and Ng \(2011\)](#) can be applied. Moreover, local identification holds for generic  $i$ , and therefore holds for any  $\theta_0$  in  $\Theta_I$  as **LCI** guarantees a unique map from  $\mathcal{V}_i$  to  $\theta_0$ .  $\square$

*Proof. of Proposition 4 ( $n_y = r$ )*

- (1) For every observable  $Y_t$  we can construct  $n_\theta$  moment conditions. Then,  $r$  is the minimum number of observables such that  $J(\theta) \equiv \frac{\partial}{\partial \theta} \Lambda(\theta)$ , is of full column rank. Thus, for  $\mathcal{H} \equiv J_{r \times n_\theta}(\theta)^T$ ,  $Z_t \in \mathbf{Y}_{t-1}$  and  $m(Z_t, \theta) \equiv \mathcal{H} \otimes \phi(Z_t)$ , the first order condition becomes  $\mathbb{E}(m(Z_t, \theta) - V(Z_t)\phi(Z_t)) = 0$  or more compactly

$\mathbb{E}m(Z_t, \theta) \equiv \mathcal{G}(\theta) = \frac{U}{r \times 1} \in (-\infty, \bar{U}]$ . Given Lemma 4,  $\mathcal{G}(\theta)$  is well behaved, and since  $n_y = r$ , it is invertible in  $U$ , which guarantees a unique solution  $\theta^*(U)$  for all admissible  $U$ . Thus,  $\Theta_I = \Theta \cap \theta^*(U)$ .

(2) Given that  $\Theta_I = \Theta \cap \mathcal{G}^{-1}(U)$ , sharpness of  $\Theta_I$  is determined by the set of admissible structures  $S$  such that  $\forall s \in S, F_{(Y_t, Y_{t-1}, Z_t)}^s : U \equiv \int U_{s,t} dF^s \in (-\infty, 0]$ .

(a) Existence of joint distribution for some  $s$  : With no frictions, that is, for  $\lambda_s = 0, \mathbb{P}_{Y_{t-1}} a.s.$ , a joint distribution for  $(Y_t, Y_{t-1}, Z_t)$  exists by construction (parametric benchmark model) and thus  $F_{(Y_t, Y_{t-1}, Z_t)}$  is equal to  $F_{(Y_t|Y_{t-1}, Z_t)} F_{Y_{t-1}} F_{Z_t}$ . Given additive separability of  $\lambda_s(Y_{t-1}, Z_t)$ ,  $F_{(Y_t, Y_{t-1}, Z_t)}^s$  also exists as  $F_{(Y_t, Y_{t-1}, Z_t)}^s = F_{(Y_t|Y_{t-1}, Z_t)}(v - \lambda_s(Y_{t-1}, Z_t)) F_{Y_{t-1}} F_{Z_t}$ .

(b) The set of admissible joint distributions: Since a joint distribution exists, the set of admissible structures are then determined by:  $U_{s,t} = \lambda_s(Y_{t-1}, Z_t) + CZ_t$  for any constant matrix  $C$ . Under Correct specification,  $U \neq 0$  iff there are frictions in the economy. This is true for the set of admissible densities  $\mathbf{F}^{s,C} \equiv \left\{ F_{(Y_t, Y_{t-1}, Z_t)}^s(v) : \int U_{s,t} dF^s < 0 \right\} \supset \mathbf{F}^{s,\lambda} \equiv \left\{ F_{(Y_t, Y_{t-1}, Z_t)}^s : \int \lambda_s(Y_{t-1}, Z_t) dF^s < 0 \right\}$ . But,  $U = 0, \forall F_{(Y_t, Y_{t-1}, Z_t)}^s \in \mathbf{F}^{s,C} \setminus \mathbf{F}^{s,\lambda}$ . The set of admissible  $U$  is the smallest set compatible with  $\lambda_s \neq 0$  and thus  $\Theta_I$  is sharp.

□

### *Proof. of Proposition 5*

This is a generalization of the proof in Proposition 5. We denote the moment conditions using  $n_y$  macroeconomic variables by  $q_1(\theta, \mathbf{Y}_{t-1})$  and moment conditions using

$m - n_y$  survey variables by  $q_2(\theta, \mathbf{Y}_{t-1})$ :

$$\begin{aligned} q_1(\theta, \mathbf{Y}_{t-1}) &= \mathcal{V}_1(\mathbf{Y}_{t-1}) \in [\underline{\mathcal{V}}(\mathbf{Y}_t)_1, \bar{\mathcal{V}}(\mathbf{Y}_t)_1] \\ &\quad \substack{n_y \times 1} \\ q_2(\theta, \mathbf{Y}_{t-1}) &= \mathcal{V}_2(\mathbf{Y}_{t-1}) \in [\underline{\mathcal{V}}(\mathbf{Y}_t)_2, \bar{\mathcal{V}}(\mathbf{Y}_t)_2] \\ &\quad \substack{(n_y - k) \times 1} \end{aligned}$$

which again imply the following unconditional moment restrictions:

$$\begin{aligned} \mathbb{E}(\phi(Z_t)(q_1(\theta, \mathbf{Y}_{t-1}) - \mathcal{V}_1(\mathbf{Y}_{t-1}))) &= 0 \\ \mathbb{E}(\phi(Z_t)(q_2(\theta, \mathbf{Y}_{t-1}) - \mathcal{V}_2(\mathbf{Y}_{t-1}))) &= 0 \end{aligned}$$

Recall that  $r$  is the minimum number of observables such that  $J(\theta) \equiv \frac{\partial}{\partial \theta} \Lambda(\theta)$ , is of full column rank. Rewrite the moment conditions such that the first  $n_\theta$  rows satisfy the rank condition. Given the total number of moment conditions used,  $m$ , the rest of the system has  $m - n_\theta$  equations. We partition  $J(\theta) \equiv (J_r(\theta), J_{n_y - r}(\theta))$  and let  $\mathcal{H}$  be an  $m \times n_\theta$  matrix where  $\mathcal{H} \equiv (J_{r \times n_\theta}(\theta)^T, J_{(n_y - r) \times n_\theta}(\theta)^T, \Delta_{\dim q_2 \times n_\theta}(\theta)^T)^T$ . Let  $m \equiv \mathcal{H} \otimes \phi(Z_t)$  and partition  $m \equiv (m_\alpha^T, m_\beta^T)^T$  where  $m_\alpha$  contains the first  $r$  elements. . Since  $m > n_\theta$ , and given a general weighting matrix  $W$ , we have the following first order condition:

$$\begin{aligned} \mathbb{E}(W_\alpha^{\frac{1}{2}T} m_\alpha + W_\beta^{\frac{1}{2}T} m_\beta - V(Z_t)\phi(Z_t)) &= 0 \\ \mathbb{E}(W_\alpha^{\frac{1}{2}T} m_\alpha + W_\beta^{\frac{1}{2}T} m_\beta - \underset{n_\theta \times 1}{U}) &= 0 \end{aligned}$$

This is a projection of  $m$  on a lower dimensional subspace. Since  $W$  is an arbitrary matrix, and  $(m_\alpha, m_\beta)$  are possibly correlated, we reproject the sum onto the space spanned by  $W_\alpha^{\frac{1}{2}T} \mathbf{m}_\alpha$ . Define  $Q_\alpha := W_\alpha^{\frac{1}{2}T} \mathbf{m}_\alpha (\mathbf{m}_\alpha^T W_\alpha^T \mathbf{m}_\alpha)^{-1} \mathbf{m}_\alpha^T W_\alpha^{\frac{1}{2}}$ , the projection, and  $Q_\alpha^\perp$  the orthogonal projection. Since the original sum satisfies the moment condition, then the two orthogonal complements will also satisfy it: Therefore,

$$\begin{aligned}
Q_\alpha \left( W_\alpha^{\frac{1}{2}T} \mathbf{m}_\alpha + W_\beta^{\frac{1}{2}T} \mathbf{m}_\beta - U \right) &= W_\alpha^{\frac{1}{2}T} \mathbf{m}_\alpha + Q_\alpha \left( W_\beta^{\frac{1}{2}T} \mathbf{m}_\beta - U \right) = 0 \\
Q_\alpha^\perp \left( W_\alpha^{\frac{1}{2}T} \mathbf{m}_\alpha + W_\beta^{\frac{1}{2}T} \mathbf{m}_\beta - U \right) &= Q_\alpha^\perp \left( W_\beta^{\frac{1}{2}T} \mathbf{m}_\beta - U \right) = 0
\end{aligned}$$

where  $U \in [W_\alpha^{\frac{1}{2}T} \underline{\mathcal{V}}(\mathbf{Y}_t)_\alpha + W_\beta^{\frac{1}{2}T} \underline{\mathcal{V}}(\mathbf{Y}_t)_\beta, W_\alpha^{\frac{1}{2}T} \bar{\mathcal{V}}(\mathbf{Y}_t)_\alpha + W_\beta^{\frac{1}{2}T} \bar{\mathcal{V}}(\mathbf{Y}_t)_\beta] \otimes \phi(\mathbf{Y}_t)$ .

As in Proposition 4, the first set of restrictions identifies a one to one mapping from  $U$  to  $\Theta_I$ , and therefore  $\theta^*(U) \equiv \mathcal{G}^{-1}(U)$ . Plugging this in the second set of restrictions eliminates dependence on  $m_\alpha$  and imposes further restrictions on the domain of variation of  $U$ . Thus

$$Q_\alpha^\perp \left( W_\beta^{\frac{1}{2}T} \mathbf{m}_\beta(U) - U \right) = 0$$

The admissible set for  $U$  is now  $\left\{ U \in (-\inf, \bar{U}) : Q_\alpha^\perp \left( W_\beta^{\frac{1}{2}T} \mathbf{m}_\beta(U) - U \right) = 0 \right\}$ . Therefore,  $\exists \theta \in \theta(U) : \theta \notin \Theta_I(U')$  and consequently  $\Theta'_I \subset \Theta_I$ .

The case of non-linear moment conditions: The above proof carries through if we replace the linear projection with conditional expectations. In this case,  $Q_\alpha$  is the conditional expectation operator, which also implies that any integrable moment function  $m$  can be decomposed as  $m = Q_\alpha m + Q_\alpha^\perp m$  such that  $Q_\alpha(m - Q_\alpha m) = 0$ . This is important for the corollary regarding survey based conditions as conditional expectations are no longer linear.  $\square$

*Proof. of Proposition 6:* Recall that agent  $i$  has the following behavioral equation  $G(\theta_1, \theta_2)x_{i,t} = F(\theta_1, \theta_2)\mathcal{E}_{t,i}(x_{i,t+1}|\mathcal{F}_{t-1,i}) + L(\theta_1, \theta_2)z_{i,t}$ , which can be rewritten as :

$$(32) \quad G(\theta_1, 0)x_{i,t} = F(\theta_1, 0)\mathbb{E}_t(x_{i,t+1}|\mathcal{F}_{t-1}) + L(\theta_1, 0)z_{i,t} + \mu_{i,t}$$

Before formally deriving the bounds, we need to establish some facts which will be used in the derivations:

(1)  $\mu_t$  is a continuous function of the state variables

The model with frictions reduces to the frictionless model by either by setting  $\theta_2 = 0$  or by eliminating differences in subjective expectations from rational expectations,  $\mathcal{E}_{i,t}(X_t|\mathcal{F}_{i,t}), \mathbb{E}(X_t|\mathcal{F}_{t,i}\forall i)$ . We show that the latter is equivalent to a change in a component of the model. Since,  $\mathcal{P}(\cdot|\mathcal{F}_{t,i})$  is absolutely continuous to the Rational Expectations measure  $\mathbb{P}(\cdot|\mathcal{F}_{t,i}\forall i)$ , there exists a Radon Nikodym derivative  $\mathcal{M}_{t,i} := \frac{d\mathcal{P}_{i,t}}{d\mathbb{P}_t}$  such that  $\mathcal{P}_{i,t}(\chi) = \int_{\chi} \mathcal{M}_{t,i} \mathbb{P}_t$  for  $\chi \subset X$ . For  $\mathcal{M}_{t,i} = 1$  for all  $t, i$  we get the frictionless model. Then given  $\theta_2$  and  $\mathcal{M}_{t,i}$ ,  $\mu_t$  is a continuous w.r.t state variables.

(2) Any probability statement about  $\mu_t$  translates to a probability statement on  $\lambda_t$ .

Since  $x_{i,t}^*$  solves the behavioral functional equation of the agent uniquely, there is a map  $h : (G, F, L) \rightarrow (P, Q)$  which is a continuous bijection, and by the implicit function theorem, any perturbation to the first order conditions (change in  $(G, F, L)$ ) maps to perturbations of the solution,  $(P, Q)$ . Therefore, for every univariate decision variable,  $\mathbb{P}(\mu_{i,t} \in [\underline{\mu}_i, \bar{\mu}_i]) = \mathbb{P}(h(\lambda_{i,t}) \in [\underline{\mu}_i, \bar{\mu}_i]) = \mathbb{P}(\lambda_{i,t} \in [\underline{\lambda}_i, \bar{\lambda}_i])$ . Given Proposition 1, same statement holds also for  $\mathbb{E}_t \lambda_{i,t}$  and  $\mathbb{E}_t \mu_{i,t}$ .

(3) Any probability statement on the subjective conditional expectations translates to a probability statement on  $\mu_{i,t}$

Recall that the enlarged state vector contains also past states,  $X_t \equiv (X_t, \tilde{X}_t)$  where  $\tilde{X}_t = X_{t-1}$ . Observing  $\mathcal{E}_{i,t}(x_{i,t+1}) \in B_t$  therefore implies observing either an expectation about the future or the present. Given the behavioral equation of the agent,  $\mathcal{E}_{i,t}(x_{i,t+1})$  maps deterministically to  $x_{i,t}$  and therefore  $\mu_{i,t}$ .

Given the above, we consider the expected value of the statistic  $\hat{\mathcal{B}}_{k,t}$ . Given the (joint) measure  $\mathbb{P} = \mathbb{P}(t) \times \Lambda(i)$ , taking the expectation we have that

$$\begin{aligned}
\mathbb{E}_t \hat{\mathcal{B}}_{k,t} &= \sum_{i \leq N} w_i \int \int 1(\mathcal{E}_{i,t}(x_{i,t+1} | \mathcal{F}_{t-1,i}) \in B) d(P(\mathcal{F}_t | \mathcal{F}_{t-1}) \times \Lambda_t(i)) \\
&= \sum_{i \leq N} w_i \bar{\mathbb{P}}_t(\mathcal{E}_{i,t}(x_{i,t+1} | \mathcal{F}_{t-1,i}) \in B) \\
&= \bar{\mathbb{P}}_t(\mathcal{E}_{i,t}(x_{i,t+1} | \mathcal{F}_{t-1,i}) \in B) \\
&= \bar{\mathbb{P}}_t(\mathbb{E}_t \mu_{i,t} \in [\mathbb{E}_t \underline{\mu}_{i,t}, \mathbb{E}_t \bar{\mu}_{i,t}]) \\
&= \bar{\mathbb{P}}_t(\mathbb{E}_t \lambda_{i,t} \in [\mathbb{E}_t \underline{\lambda}_{i,t}, \mathbb{E}_t \bar{\lambda}_{i,t}]) = B_t
\end{aligned}$$

In the second to last equality, we use fact **3** and in the last equality we have used fact **2** adapted to conditional means. We next derive the main result.

Let  $Y_t^o = X_t$ ,  $Y_t^m = P(\theta_1, 0)X_{t-1}$ . Denote by  $(\mathcal{X}_t^*, \mathcal{Z}_t^*)$  the time dependent subset of the support such that  $\lambda_{i,t} < 0, \forall (x_{i,t-1}, z_{i,t}) \in (\mathcal{X}_t^*, \mathcal{Z}_t^*)$  and by  $(\mathcal{X}_t^{*,C}, \mathcal{Z}_t^{*,C})$  their complements in  $(\mathcal{X}, \mathcal{Z})$ . Given that  $\lambda_i^{1,2}$  can possibly differ across  $i$ , let<sup>24</sup>

$$(33) \quad \lambda_{i,t}^j \equiv \begin{cases} \lambda^1 x_{i,t-1} \text{ or } \lambda^2 z_{i,t} & \text{if } (x_{i,t-1}, z_{i,t}) \in (\mathcal{X}_t^*, \mathcal{Z}_t^*) \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned}
Y_t^o &= \int \lambda_t^1 d\Lambda_t(i) + \int \lambda_t^2 d\Lambda_t(i) + Q(\theta_1, 0)Z_t + Y_t^m \\
&= \lambda^1 \int_{\mathcal{X}_t^*} x_{i,t-1} d\Lambda_t(i) + \lambda^2 \int_{\mathcal{Z}_t^*} z_{i,t} d\Lambda_t(i) + Q(\theta_1, 0)Z_t + Y_t^m
\end{aligned}$$

<sup>24</sup>The typical example in this case would be economies with ex post heterogeneity i.e. across net asset positions.

Pre-multiplying and post dividing every element of  $\int_{\mathcal{X}_t^*} x_{i,t-1} d\Lambda_t(i)$  and  $\int_{\mathcal{Z}_t^*} z_{i,t} d\Lambda_t(i)$  with the corresponding aggregates:

$$(34) Y_t^o = \lambda^1 X_{t-1} \odot \int_{\mathcal{X}_t^*} x_{i,t-1} d\Lambda_t(i) \oslash X_{t-1} + \lambda^2 \odot Z_t \int_{\mathcal{Z}_t^*} z_{i,t} d\Lambda_t(i) \oslash Z_t + Q(\theta_1, 0) Z_t + Y_t^m$$

where  $\odot$  and  $\oslash$  are the Hadamard product and division respectively, and  $\mathbb{P}$  is the probability over vectors  $z_{i,t}$  and  $x_{i,t-1}$  using  $\Lambda_i$ . Without loss of generality, looking at the first element of the vector  $\int_{\mathcal{X}_t^*} x_{i,t-1} d\Lambda_t(i) \oslash X_{t-1}$ ,

$$\begin{aligned} & \int_{\mathcal{X}_t^*} x_{i,t-1}^1 d\Lambda_t(i) \oslash X_{t-1}^1 \\ &= \frac{\mathbb{E}(x_{i,t-1}^1 | x_{i,t} \in \mathcal{X}_t^*) \mathbb{P}(x_{i,t} \in \mathcal{X}_t^*)}{\mathbb{E}(x_{i,t-1}^1 | x_{i,t} \in \mathcal{X}_t^*) \mathbb{P}(x_{i,t} \in \mathcal{X}_t^*) + \mathbb{E}(x_{i,t-1}^1 | x_{i,t} \in \mathcal{X}_t^{*,C}) \mathbb{P}(x_{i,t} \in \mathcal{X}_t^{*,C})} \\ &= \frac{\mathbb{P}(x_{i,t} \in \mathcal{X}_t^*)}{\mathbb{P}(x_{i,t} \in \mathcal{X}_t^*) + \mathbb{P}(x_{i,t} \in \mathcal{X}_t^{*,C}) + \left( \frac{\mathbb{E}(x_{i,t-1}^1 | x_{i,t} \in \mathcal{X}_t^{*,C})}{\mathbb{E}(x_{i,t-1}^1 | x_{i,t} \in \mathcal{X}_t^*)} - 1 \right) \mathbb{P}(x_{i,t} \in \mathcal{X}_t^{*,C})} \\ &:= \frac{\mathbb{P}(x_{i,t} \in \mathcal{X}_t^*)}{1 + \hat{R}_t^1} = \frac{B_t}{1 + \hat{R}_t^1} \leq \frac{B_t}{1 + \inf_{\mathcal{X}_t^*} \hat{R}_t^1} = B_t \end{aligned}$$

The last inequality holds since  $\inf_{\mathcal{X}_t^*} \hat{R}_t^1 = 0$ ,  $\sup_{\mathcal{X}_t^*} \hat{R}_t^1 = \infty$  and by the monotonicity assumption **(S-4)**. The inequality holds for every element of  $x_{i,t}$ . Finally, taking conditional expectations of (34) (using  $\mathbb{P}(\cdot | X_{t-1})$ ):

$$\mathbb{E}(Y_t^o | X_{t-1}) \leq \lambda^1 X_{t-1} \odot \mathbb{E}(B_t | X_{t-1}) + Y_t^m$$

□

### *Proof. of Corollary 7*

(1): When  $B_t = 0$ ,  $\mathbb{P}(x_{i,t} \in \mathcal{X}_t^*) = 0$  and therefore the model has no frictions, which trivially restores point identification. When  $B_t = 1$ ,  $\int_{\mathcal{X}_t^*} x_{i,t-1}^1 d\Lambda_t(i) = X_{t-1}$ , and this results to the first type of moment inequality using macroeconomic data. We now show the possibility of refinement of  $\Theta$  for  $B \in (0, 1)$ . Defining the conditional errors from the first and second types of inequalities as  $\nu_t \equiv Y_t^o - P(\theta_1, 0) X_{t-1}$  and  $\eta_t \equiv \hat{Y}_t^o - P(\theta_1, 0) X_{t-1} - \lambda^1 X_{t-1} \odot \mathbb{E}(B_t | X_{t-1})$  respectively,  $\mathbb{E}(\eta_t | \nu_t, \mathbf{Y}) = \int \eta_t p(\eta_t | \nu_t, \mathbf{Y}) \neq \eta_t$  since  $p(\eta_t | \nu_t, \mathbf{Y}) \neq 1$  and  $\eta_t - \mathbb{E}(\eta_t | \nu_t, \mathbf{Y}) \neq 0$  unless  $B_t = 0$ .

(2): Suppose that  $\Theta_I$  is a singleton. This implies that  $B = 0$ , or  $\Lambda_t(i)$  has unit mass on one agent which is unconstrained, and thus  $B = 0$  too.  $\square$

**Proof. of Proposition 8**

Consider the quantity  $T^{-\frac{1}{2}} \inf_{\theta \in \hat{\Theta}_s} \|\mathcal{V}^{-\frac{1}{2}}(\hat{f}(\hat{\theta}_p) - \hat{f}(\hat{\theta}_s))\|$ . Given that  $\Theta$  is a connected set and  $\hat{\Theta}_s \in \Theta$ , then  $\hat{\Theta}_s$  is also connected. For any  $\theta, \theta' \in cl(\hat{\Theta}_s)$ ,  $d(\theta, \theta') < \epsilon$  for arbitrarily small  $\epsilon > 0$ . This implies that if  $\theta_p \in \Theta_s$  then there exists a  $\theta_s \in \hat{\Theta}_s$  such that  $\|\mathcal{V}^{-\frac{1}{2}}(f(\theta_p) - f(\theta_s))\| < \epsilon$ . For every estimating equation, redefine  $\hat{f}(\hat{\theta}_p) = T^{-1} \sum_t q(Y; \theta) - T^{-1} \sum_t \mathbb{E}q(Y; \theta) + \gamma_T = A_T(\theta) + \gamma_T$  where  $A_T(\theta) \equiv (A_{T,1}(\theta), A_{T,2}(\theta), \dots, A_{T,p}(\theta))^T$  and  $\gamma_T = (\gamma_{1,T}, \gamma_{2,T}, \dots, \gamma_{p,T})^T$ . By element-wise mean value expansion around  $\theta_l \in \Theta_l, l \in s, p$ ,  $\hat{f}(\hat{\theta}) = \mathbb{E}q(Y_t, \theta_l) + (\hat{f}(\theta_l) - \mathbb{E}q(Y_t, \theta_l)) + D(\tilde{\theta})(\theta - \theta_l)$ . Given mild assumptions on the  $(2 + \delta)$  boundedness of each moment, the second component scaled by  $\sqrt{T}$ ,  $\sqrt{T}(\hat{f}(\theta_l) - \mathbb{E}\hat{f}(\theta_l)) \rightarrow_d \mathcal{N}(0, \Omega_{1,l})$  while the scaled third component,  $D(\tilde{\theta})\sqrt{T}(\theta - \theta_l) \rightarrow_d \mathcal{N}(0, \Omega_{2,l})$ . Consequently, for each  $j, j \leq p$ ,  $A_{T,j} \rightarrow_d \mathcal{N}(0, \Omega_j^*)$ . Let  $\mathcal{V}(\theta) = \text{AsyVar}(T^{\frac{1}{2}}f(\theta_p) - T^{\frac{1}{2}}f(\theta_s))$  and denote by  $\mathcal{V}_d(\theta)$  be the matrix containing only the diagonal elements of  $\mathcal{V}(\theta)$ .

$$T^{\frac{1}{2}}\mathcal{V}_d^{-\frac{1}{2}}(\hat{f}(\hat{\theta}_p) - \hat{f}(\hat{\theta}_s)) = \mathcal{V}_d^{-\frac{1}{2}}T^{\frac{1}{2}}(A_T(\theta_p) - A_T(\theta_s) + \gamma_{T,p} - \gamma_{T,s})$$

Under  $H_0$ ,  $\theta_p \in \Theta_s$ , or  $f(\theta_p) \in f(\Theta_s)$  and thus  $\inf_{\theta \in \Theta_s} (\gamma_{T,p} - \gamma_{T,s}) = 0$

$$\begin{aligned} TW(\theta_p, \Theta_s) &= \inf_{f(\theta_s) \in f(\hat{\Theta}_s)} \|\hat{\mathcal{V}}_d^{-\frac{1}{2}}(T^{\frac{1}{2}}A_T(\theta_p) - T^{\frac{1}{2}}A_T(\theta_s) + T^{\frac{1}{2}}(\gamma_{T,p} - \gamma_{T,s}))\|^2 \\ &= \inf_{\theta_s \in \hat{\Theta}_s} \|\hat{\mathcal{V}}_d^{-\frac{1}{2}}(T^{\frac{1}{2}}A_T(\theta_p) - T^{\frac{1}{2}}A_T(\theta_s) + T^{\frac{1}{2}}(\gamma_{T,p} - \gamma_{T,s}))\|^2 \\ &\xrightarrow{d} \|\mathcal{V}_d^{-\frac{1}{2}}\mathcal{V}^{\frac{1}{2}}\mathcal{N}(0, I_p)\|^2 \end{aligned}$$

Under  $H_1$ ,  $\theta_p \neq \Theta_s$ , and therefore  $\inf_{\theta_s \in \hat{\Theta}_s} (\gamma_{T,p} - \gamma_{T,s}) = O(1)$ .

$$\begin{aligned}
T\mathcal{W}(\theta_p, \Theta_s) &= \inf_{f(\theta_s) \in f(\hat{\Theta}_s)} \|\hat{\mathcal{V}}^{-\frac{1}{2}} T^{\frac{1}{2}} (A_T(\theta_p) - A_T(\theta_s) + \gamma_{T,p} - \gamma_{T,s})\|^2 \\
&= \|\hat{\mathcal{V}}^{-\frac{1}{2}} (T^{\frac{1}{2}} A_T(\theta_p) - T^{\frac{1}{2}} A_T(\theta_s) + \inf_{f(\theta_s) \in f(\hat{\Theta}_s)} T^{\frac{1}{2}} (\gamma_{T,p} - \gamma_{T,s}))\|^2 \\
&= \|O_p(1) + O_p(T^{\frac{1}{2}})\|^2 = O_p(T)
\end{aligned}$$

□

*Proof. of  $\mu_{ols} \notin \mu_{ID,2}$  in Partial Equilibrium*

We first solve for the consumption process. Let  $p_t := \mathbb{P}_t(\lambda_{i,t} > 0)$  and  $p := \frac{\mathbb{E}(c_{i,t}^2 \mathbb{P}_t(\lambda_{i,t} > 0))}{\mathbb{E}c_{i,t}^2}$ . Moreover, let  $\beta(1+r)$  and  $\rho \rightarrow \infty$ ,  $\beta(1+r)^{\frac{1}{\rho}} = 1$  and therefore  $\tilde{\mu}_0 = 0$ . Denote by  $\varphi(u)$  the Normal density at  $u$  and by  $\mathfrak{s}(\cdot)$  the marginal density of  $c_{i,t}$ , Assuming that  $c_{i,t}$  started  $\tau$  periods ago,

$$c_{i,t+1} = \sum_{j=0}^{\tau} \epsilon_{t+1-j} (1 + \chi_{i,t-j} \lambda_2) \prod_{k=0}^{j-1} (1 + \chi_{i,t-k} \lambda_1) := \sum_{j=0}^{\tau} \epsilon_{t+1-j} z_{t+1-j} := \sum_{j=0}^{\tau} \xi_{t+1-j}$$

where  $\mathbb{P}(z_{t+1-j} = 0) = 0$ . Consumption is a convolution of  $\tau+1$  independent variables  $\xi_{t+1-j}$  whose marginal density is a mixture of  $N(0, \omega(\lambda_1, \lambda_2, j))$  with weights that depend on  $p$  and  $j$ . For example,  $\xi_{t+1-j} = \epsilon_{t+1-j} (1 + \lambda_2) \prod_{k=0}^{j-1} (1 + \lambda_1)$  w.p.  $(1-p)^{j+1}$ . For each  $\xi_{t+1-j}$ , symmetry around zero is maintained. Next, we prove the result for  $\tau = 1$ . Evaluating the LHS of (7) at the true  $\tilde{\mu}_0$  and the RHS at  $\mu_{ols}$ , setting  $u = 0$  and taking their difference,

$$\Delta := \left(\frac{1}{2} - \Phi_{0, \sigma_\epsilon^2}(-(\tilde{\mu}_{ols} c_{i,t}))\right)(1 - p_t) + (\Phi_{0, \sigma_\epsilon^2(1+\lambda_2)^2}(-(\tilde{\mu} + \lambda_1) c_{i,t}) - \Phi_{0, \sigma_{ols}^2}(-\tilde{\mu}_{ols} c_{i,t})) p_t$$

For  $\mu_{ols}$  to be admissible,  $\Delta$  should be weakly positive for all  $p_t$  such that

$(\Phi_{0, \sigma_\epsilon^2(1+\lambda_2)^2}(-(\tilde{\mu} + \lambda_1) c_{i,t}) - \Phi_{0, \sigma_{ols}^2}(-\tilde{\mu}_{ols} c_{i,t})) \geq 0$ . Sufficient conditions for the latter, are  $\sigma_{ols}^2 = \sigma_\epsilon^2(1 + \lambda_2)^2$ , which is true for all  $(\lambda_1, \lambda_2) : \lambda^2 + 2\lambda_2 - \lambda_1 \frac{\sigma_\epsilon^2}{\sigma_{ols}^2} p(1-p) = 0$ , and for probability functions  $B_t : \lambda_1 c_{i,t} (1-p) \geq 0$ . We choose  $B_t = 1(c_{i,t} < 0)$ .

Substituting for the test point  $\tilde{\mu} = \mu_{ols}$  and  $B_t = 1(c_{i,t} < 0)$ ,  $\Delta > 0$  for  $c_{i,t} > 0$  by construction, while for  $c_{i,t} < 0$  we have a contradiction:

$$(35) \quad \Delta = \frac{1}{2} - \Phi_{0,\sigma_\epsilon^2}(-p\lambda_1 c_{i,t}) = \frac{1}{2} - \Phi_{0,\sigma_\epsilon^2}(p\lambda_1 |c_{i,t}|) < 0$$

□

10.0.1. *Small Open Economy model for Spain.* In what follows variables with \* denote the "rest of the world",  $y_t$  is real output,  $c_t$  is consumption,  $i_t$  investment,  $q_t$  the value of capital,  $k_t$  is productive capital,  $k_t^s$  capital services,  $z_t$  is capital utilization,  $\mu_t^p$  is the price markup,  $\pi_t$  is domestic inflation,  $\pi_{cpi}$  is CPI inflation,  $r_t^k$  is the rental rate of capital,  $w_t$  is the real wage and  $r_t$  is the interest rate

$$\begin{aligned}
y_t &= c_y c_t + i_y i_t + z_y z_t + n x_y s_t + \epsilon_t^g \\
c_t &= c_1 \mathbb{E}_t c_{t+1} + c_2 (l_t - \mathbb{E}_t l_{t+1}) - c_3 (r_t - \mathbb{E}_t \pi_{t+1} + \epsilon_t^b) \\
\mathbb{E}_t r_{t+1}^k &= r_1 (r_t - \mathbb{E}_t \pi_{t+1} + \epsilon_t^b) \\
y_t &= \phi_p (\alpha k_t^s + (1 - \alpha) l_t + \epsilon_t^\alpha) \\
c_t &= c_t^* + \frac{1}{\sigma_a} s_t \\
k_t^s &= k_{t-1} + z_t \\
z_t &= z_1 r_t^k \\
k_t &= k_1 k_{t-1} + (1 - k_1) i_t \\
\mu_t^p &= \alpha (k_t^s - l_t) + \epsilon_t^\alpha - w_t \\
\pi_t &= \pi_1 \pi_{t-1} + \pi_2 \mathbb{E}_t \pi_{t+1} - \epsilon_t^p \\
\pi_{cpi} &= \pi_t + \nu \Delta s_t \\
r_t^k &= -(k_t - l_t) + w_t \\
\mu_t^w &= w_t - \sigma_l l_t + c_t \\
w_t &= w_1 w_{t-1} + (1 - w_1) (\mathbb{E}_t w_{t+1} + \mathbb{E}_t \pi_{t+1}) - w_2 \pi_t + w_3 \pi_{t-1} - w_4 \mu_t^w + \epsilon_t^w \\
r_t &= \rho r_{t-1} + (1 - \rho) [r_\pi \pi_{cpi,t} + r_y (y_t - y_t^p)] + r_{\Delta Y} [(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)] + \epsilon_t^r
\end{aligned}$$

Financial Frictions and Capital adjustment costs imply the following changes to the model, where  $Q_t$  is the price of capital and  $R_t^k$  the return to capital :

$$\begin{aligned}
\mathbb{E}_t R_{t+1}^k &= -\chi_{pr}(N_t - Q_t - k_t) + r_t - \mathbb{E}_t \pi_{t+1} \\
N_t &= \gamma \bar{r}_k \left( \left( \frac{\bar{K}}{\bar{N}} \right) (R_t^k - r_{t-1}) + r_t^k + N_{t-1} \right) \\
i_t &= i_1 i_{t-1} + (1 - i_1) \mathbb{E}_t i_{t+1} + \frac{1}{\phi_{adj}} Q_t + \epsilon_{q,t} \\
Q_t &= q_1 \mathbb{E}_t Q_{t+1} + (1 - q_1) \mathbb{E}_t r_{t+1}^k - (R_t^k + \epsilon_{b,t})
\end{aligned}$$

TABLE 1. Confidence Set for  $\Theta_{IM}$  and  $\Theta_{CM}$ -with survey data

Parameter	Incomplete Model		Complete Model	
	$q_{2.5\%}$	$q_{97.5\%}$	$q_{5\%}$	$q_{95\%}$
$\sigma_c$	12.460	13.320	28.6900	30.0000
$\phi_p$	5.1840	5.3430	21.5700	21.7000
$g_y$	0.9668	1.0000	0.2432	0.4827
$\lambda_h$	0.7550	0.7673	0.0009	0.0009
$\sigma_l$	8.5140	8.5760	8.892	9.104
$\xi_p$	0.3654	0.3846	0.0000	0.0000
$\xi_w$	0.0945	0.0958	0.8173	0.8628
$\iota_w$	0.5631	0.575	0.9748	0.9999
$\iota_p$	0.4952	0.498	0.0004	0.0004
$z_1$	0.5360	0.5429	0.0009	0.0009
$\rho_a$	0.6447	0.6504	0.7521	0.9092
$\rho_b$	0.6729	0.6789	<i>0.0000</i>	<i>0.0000</i>
$\rho_g$	0.0000	0.0023	0.9999	0.9999
$\rho_{qs}$	0.3394	0.3468	0.3242	0.4079
$\rho_{ms}$	0.4044	0.4137	0.7347	0.8015
$\rho_p$	0.4802	0.4843	0.6329	0.7077
$\rho_w$	0.2690	0.2719	0.0003	0.0003
$ma_p$	0.9958	1.000	0.0129	0.0561
$ma_w$	0.5466	0.5487	0.0010	0.0011
$\nu$	0.6084	0.6104	0.0053	0.0205
$\rho_y^\star$	0.6080	0.6080	0.1683	0.2251
$\sigma_a$	9.8980	20.000	0.0245	0.0339
$\sigma_b$	3.8910	7.3910	3.4480	3.4800
$\sigma_g$	13.530	13.830	4.5820	4.6340
$\sigma_{qs}$	9.2380	15.23	0.1831	0.2064
$\sigma_m$	9.1220	10.240	1.0250	1.0670
$\sigma_p$	0.9563	1.9220	4.2260	4.2660
$\sigma_w$	-	-	-	-
$\sigma_f$	-	-	-	-
$\phi_{adj}$	-	-	5.9210	5.9550
$\chi_{pr}$	-	-	0.9862	0.9999
$\gamma$	-	-	0.9944	0.9950
$K/N$	-	-	4.3330	4.3490

Note:  $\sigma_w$  and  $\sigma_f$  are set to 1. Parameters in italic font reached the boundary and thus calibrated to the boundary value.

Estimated Wedges - with Survey Data- 5% LEVEL

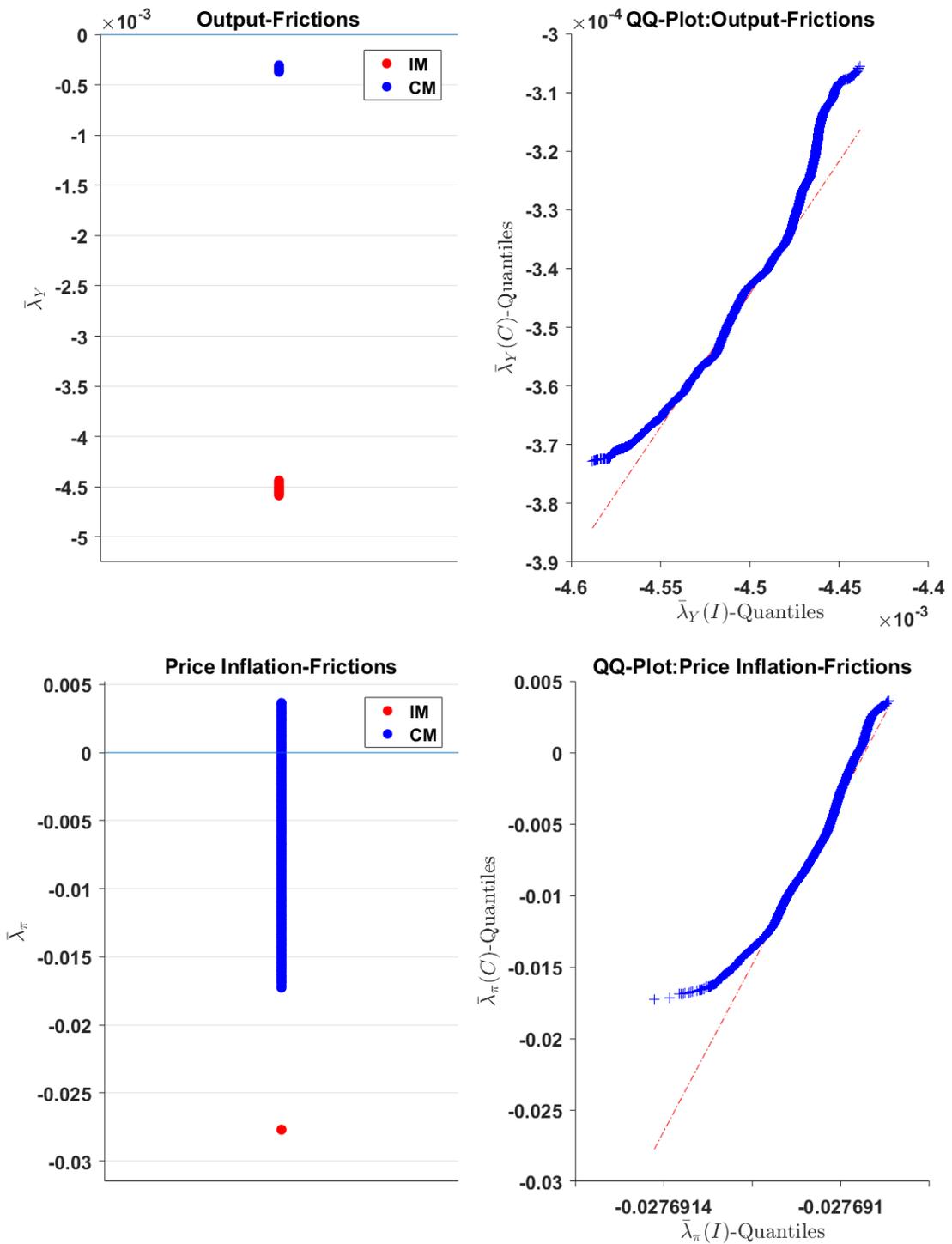


FIGURE 1. Full Sample Wedges (due to Financial Frictions and Capital Adjustment Costs)

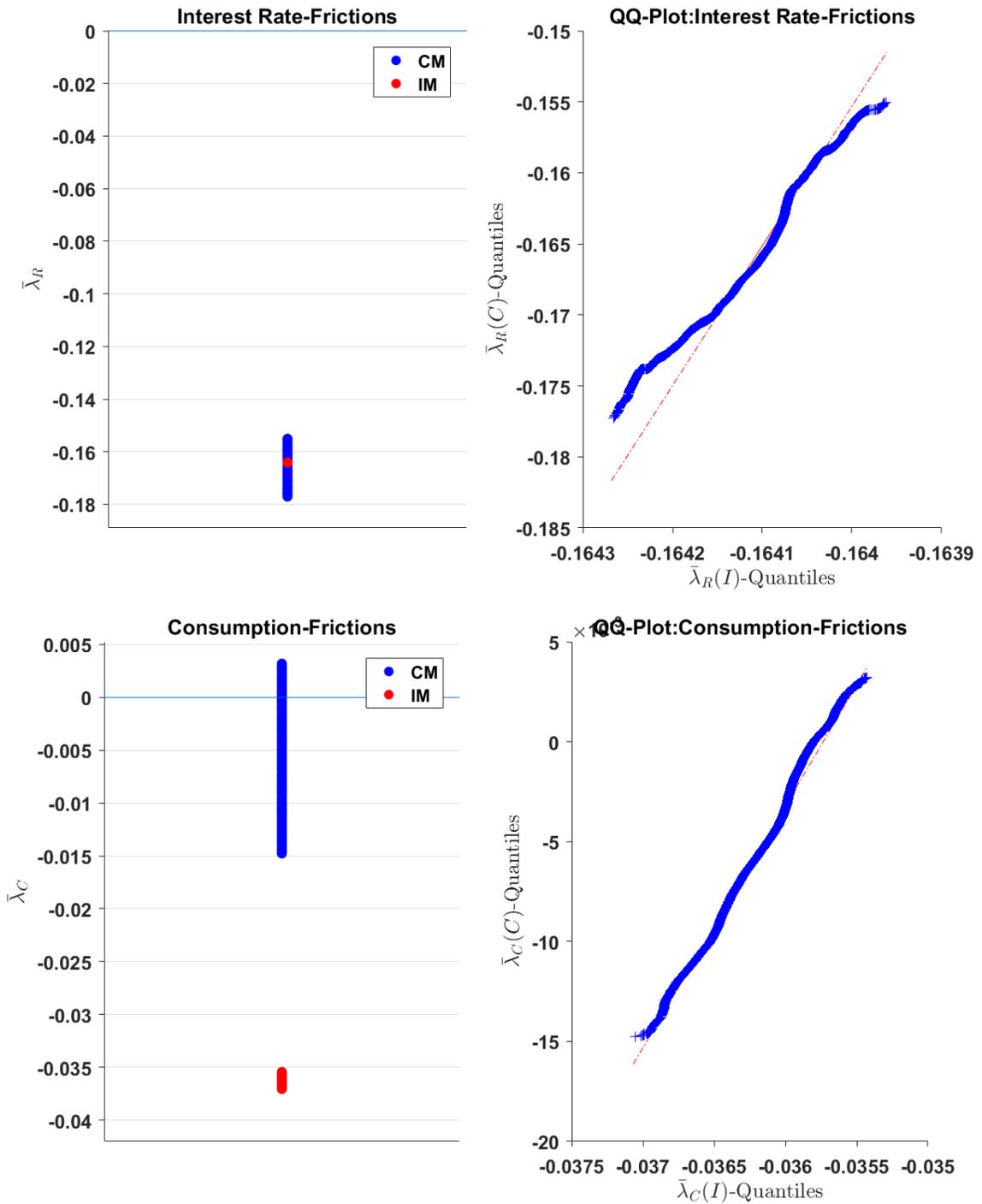


FIGURE 2. Full Sample Wedges (due to Financial Frictions and Capital Adjustment Costs)

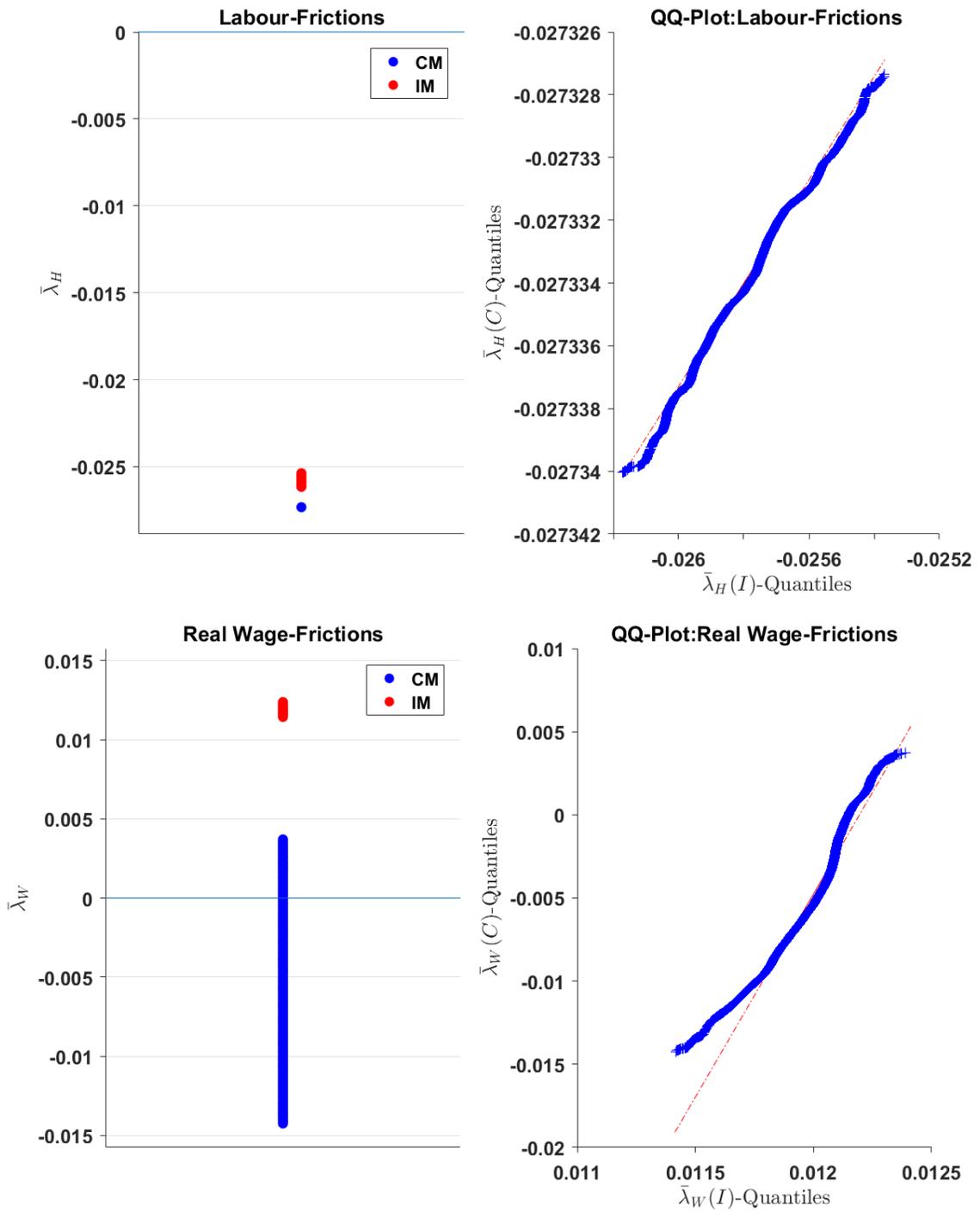


FIGURE 3. Full Sample Wedges (due to Financial Frictions and Capital Adjustment Costs)

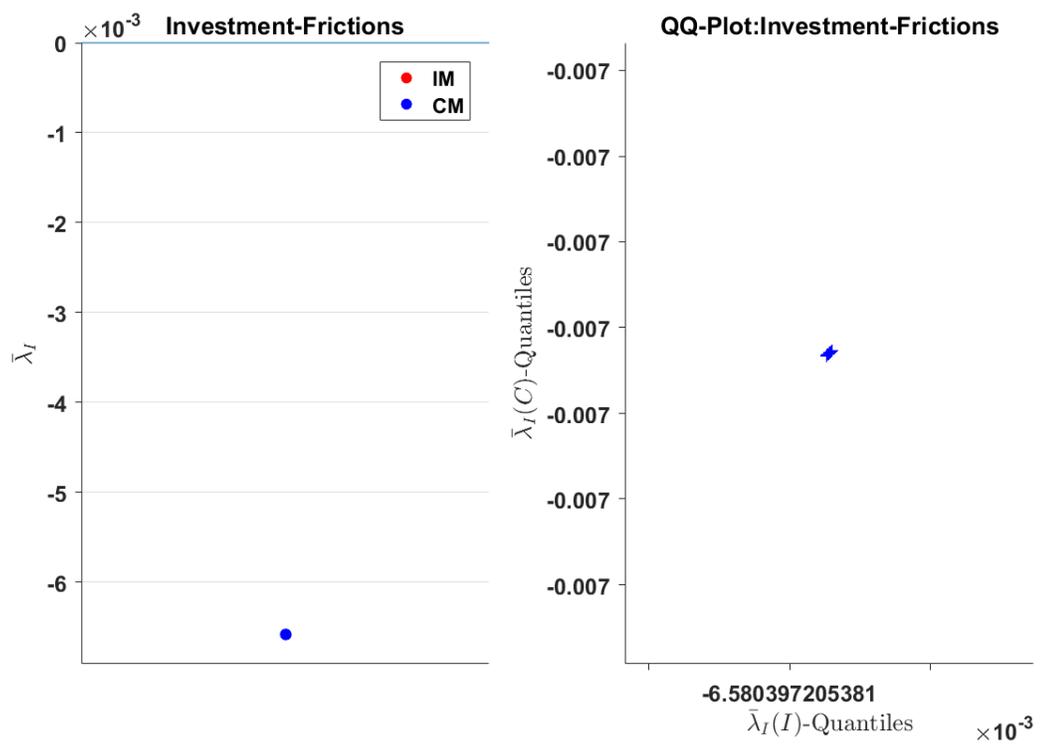


FIGURE 4. Full Sample Wedges (due to Financial Frictions and Capital Adjustment Costs)