C.I.R.P.É.E

Centre Interuniversitaire sur le Risque, les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **07-47**

Using Implied Probabilities to Improve Estimation with Unconditional Moment Restrictions

Alain Guay
Florian Pelgrin

Décembre/December 2007

Guay: Université du Québec à Montréal, CIRPÉE and CIREQ
guay.alain@uqam.ca
Pelgrin: Université de Lausanne – HEC, DEEP, CIRANO and IEMS
fpelgrin@unil.ch

A previous version of this paper has circulated under the title "The U.S. Keynesian Phillips Curve: An Empirical Assessment". Florian Pelgrin gratefully acknowledges financial support from the National Center of Competence in Research "Financial Valuation and Risk Management". The National Centers of Competence in Research (NCCR) are a research instrument of the Swiss National Science Foundation. Alain Guay gratefully acknowledges financial support from the Social Sciences and Humanities Research Council (SSHRC) and le Fonds québécois de la recherche sur la société et la culture (FQRSC). The views in this paper are those of the authors and not necessarily those of their respective institutions.

Abstract:
In this paper, we investigate the information content of implied probabilities (Back and Brown, 1993) to improve estimation in unconditional moment conditions models. We propose and evaluate two 3-step euclidian empirical likelihood estimators and their bias-correction versions for weakly dependent data. The first one is the time series extension of the 3S-EEL proposed by Antoine, Bonnal and Renault (2007). The second one is new and uses in contrast only an estimator of the weighting matrix at an efficient 2-step GMM estimator, while leaving unrestricted the Jacobian matrix. Both estimators use implied probabilities to achieve higher-order improvements relative to the traditional GMM estimator. A Monte-Carlo study reveals that the finite and large sample properties of the (bias-corrected) 3-step estimators compare very favorably to the existing approaches: the 2-step GMM and the continuous updating estimator. As an application, we re-assess the empirical evidence regarding the New Keynesian Phillips curve in the US.

Keywords: Information-based inference, Implied probabilities, Weak identification, Generalized method of moments, Phillips curve

JEL Classification: C13, C14, E31

1 Introduction

A number of studies have recently revealed that the efficient Generalized Method of Moments (GMM) introduced by Hansen (1982) may have large bias for sample size typically encountered in applied economics.[1] Alternative estimators based on a one-step procedure that are first-order equivalent to GMM have been suggested to address this problem. Newey and Smith (2004) have shown that these alternative estimators share a common structure, being members of a class of generalized empirical likelihood (GEL) estimators. These alternative estimators include the Continuous Updating Estimator (CUE) proposed by Hansen, Heaton and Yaron (1996), the Empirical Likelihood (EL) estimator of Qin and Lawless (1994), and the Exponential Tilting (ET) estimator of Kitamura and Stutzer (1997) and Imbens, Spady and Johnson (1998). Newey and Smith (2004) in an i.i.d. context and Anatolyev (2005) for weakly dependent data have shown that these one-step estimators achieve asymptotic higher-order improvements relative to the traditional 2-step GMM estimator. However, these estimators are more computationally demanding. Especially for the EL and ET estimators, this requires the optimization of a saddle point problem. Moreover, recent studies on the finite sample properties suggest that these estimators may have a larger root mean square error than the traditional 2-step GMM estimator (see Guggenber and Hahn 2005).

In this paper, we investigate the information content of implied probabilities to improve estimation in unconditional moment conditions models. As introduced by Back and Brown (1993), implied probabilities assign a weight to each observation in the sample such that moment conditions are satisfied. In particular, more (respectively less) weight is assigned to an observation for which the moment restrictions are (respectively not) satisfied at the parameter estimates. In that respect, as suggested by Back and Brown (1993), implied probabilities can then provide a useful diagnostic device for model specification. Moreover, the information content of implied probabilities can also be exploited to provide efficient moment estimators as shown by Brown and Newey (1998) in an i.i.d. context, and Smith (2004) for weakly dependent data. For instance, efficient estimators of the Jacobian and the optimal weighting matrices can be obtained by using implied probabilities instead of the uniform weights, $1/T$. Estimators based on moment conditions computed with such efficient estimators of the Jacobian and optimal weighting matrices have shown to achieve an asymptotic higher-order improvements with respect to the traditional 2S-GMM (Newey and Smith, 2004; Anatolyev, 2005).

Our objective is thus to improve the performance of the 2-step GMM estimator by using implied probabilities. At the same time, we seek to preserve its computational simplicity. On the one hand, we built on the work done by Antoine, Bonnal and Renault (2007)[2] In an i.i.d context, Antoine, Bonnal and Renault (2007) propose a 3S-EEL estimator based on a Chi-

[1] For instance, see the special number of *Journal of Business and Economic Statistics*, July, 1996.

[2] See also Bonnal and Renault (2001), and Smith (2007).

square distance where the last step consists of solving the First Order Conditions (FOC) of the EL estimator given some efficient estimators of the Jacobian and the optimal weighting matrices evaluated at an efficient estimator (usually the traditional 2S-GMM estimator). Using smoothed moment conditions, we show how to extend their estimator to time-series models. On the other hand, we propose a new 3-step estimator (called 3SW-EEL) that only uses an efficient optimal weighting matrix based on the 2S-GMM implied probabilities—the Jacobian matrix is left unrestricted. Finally, an analytical bias correction, which also uses implied probabilities, is provided for both estimators. These estimators have three appealing properties. First, in contrast to the (smoothed) Generalized Empirical Likelihood (GEL) estimator, they avoid to solve a computationally demanding saddle point problem, which grows with the number of moment conditions. Second, both estimators achieve a higher-order equivalence to the SEL (up to an order $\mathcal{O}_p(T^{-3/2})$) and their bias-corrected versions are asymptotically unbiased up to order T^{-1}. Finally, the proposed bias correction has the advantage of being computationally much simpler than the bootstrap or jackknife correction methods.

In order to evaluate our proposed estimators, we run extensive Monte Carlo simulations. More specifically, we compare the finite and large sample properties of our estimators with those of the 2S-GMM estimator and the CUE. We assume that the data generating process is given by the reduced-form of a univariate linear rational expectations model. This class of models is often used in applied macroeconomics, as for instance any log-linearized Euler equation in a dynamic stochastic general equilibrium model. Therefore, our results are of particular interest and can provide some useful guidelines in applied economics. Simulation results provide evidence that our proposed estimators are very competitive with respect to the 2S-GMM estimator and the CUE. More specifically, they almost always perform better in terms of mean bias and root mean squared error than the 2S-GMM estimator. Among the proposed smoothed 3S-EEL estimators, the 3SW-EEL estimator and its bias-corrected version have generally better finite and large sample properties than the time-series extension of the 3S-EEL estimator.

Moreover, we also provide an empirical application regarding the New Keynesian Phillips curve (NKPC) in the US. In doing so, we propose a new specification J-based test-statistic, which measures the discrepancy between the euclidian implied probabilities and the unconstrained empirical probabilities $1/T$. Moreover, as is well known by now, we take care of weak identification by proposing two test-statistics that are robust to weak identification. Overall, we find evidence that the inflation dynamics is mostly forward-looking and driven by the forcing variable—the real marginal cost. However, weak identification cannot be ruled out and misspecification is an issue with large instrument sets.

The rest of the paper is organized as follows. In Section 2, we define the concept of implied probabilities as in Back and Brown (1993). Section 3 presents the two (bias-corrected) smoothed 3-step estimators. In Section 4, we provide Monte Carlo simulations. In section 5, we re-assess the empirical evidence regarding the new Keynesian Phillips curve in the US. The last section concludes. All proofs are relegated to the Appendix.

2 Implied Probabilities

In this section, we define the concept of implied probabilities, which is used throughout the paper. We consider models specified by a finite number of moment conditions. More precisely, let $\{z_t : t = 1, \cdots, T\}$ be $\mathbb{R}^l$ -valued time series data, where T denotes the sample size. Let $g(z_t, \theta)$: $H \times \Theta \to \mathbb{R}^q$, where $H \subset \mathbb{R}^l$ and $\Theta \subset \mathbb{R}^p$, and $\theta \in \Theta$ denote respectively the parameter space and the p-vector of unknown parameters. The number of moment conditions, q, exceeds or is equal to the number of parameters, p. The true parameter vector θ_0 satisfies the unconditional moment conditions:

$$E\left[g(z_t, \theta^0)\right] = 0 \tag{1}$$

where $E[\cdot]$ denotes the expectation operator with respect to the unknown distribution of z_t.

To introduce the concept of implied probabilities, consider the optimal 2S-GMM estimator. It minimizes the following objective function over $\theta \in \Theta$

$$\hat{\theta}_T^{2S} = \arg\min_{\theta \in \Theta} \frac{1}{T} \sum_{t=1}^{T} g(z_t, \theta)' \hat{\Omega}_T(\hat{\theta}_T^{1S})^{-1} \frac{1}{T} \sum_{t=1}^{T} g(z_t, \theta) \tag{2}$$

where $\hat{\theta}_T^{1S}$ is a first-step estimator, usually obtained with the identity matrix as a weighting matrix, and $\hat{\Omega}_T^{-1}$ is an $\mathcal{O}_p(1)$ positive definite weighting matrix, which is a consistent estimator of the inverse of the variance-covariance matrix of the moments conditions.

The 2S-GMM estimator only uses information from the just-identified moment conditions. However, as pointed out by the empirical likelihood literature (Baggerly, 1988; Owen, 1990, 1991, 2001; Qin and Lawless, 1994; Smith, 2000), over-identified moment restrictions can be helpful to revise our empirical view about the DGP and to bring useful information about some characteristics of the DGP. The main idea is to seek implied probabilities that precisely afford an efficient use of the information content of estimating equations. In this respect, Back and Brown (1993) propose a distribution function estimator of the data based on the moment conditions for improving the estimation of an arbitrary measurable function. As defined by Back and Brown (1993), the corresponding (2S-GMM) implied probabilities of the distribution function are given by

$$p_t^{GMM}(\theta) = \frac{1}{T} - \frac{1}{T-p} \left[J_{tT}(\theta) - \overline{J}_T(\theta)\right]' \hat{\Omega}_T(\theta)^{-1} \overline{g}_T(\theta) \tag{3}$$

with $\overline{J}_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} J_{tT}(\theta)$ and

$$J_{tT}(\theta) = \sum_{s=1}^{T} \kappa(|t-s|) g(z_{t-s}, \theta)$$

where $\kappa(|t-s|)$ is a real valued weighting function and $\overline{g}_T(\theta) = \frac{1}{T}\sum_{l=1}^{T} g(z_t,\theta)$. It turns out that a consistent and positive definite estimator of the variance-covariance matrix is given by

$$\hat{\Omega}_T(\theta) = \frac{1}{T-p}\sum_{t=1}^{T} J_{tT}(\theta) g(z_t,\theta)'.$$

It has the usual form of a Heteroskedasticity and Autocorrelation Consistent (HAC) weighting matrix for standard kernels $\kappa(|t-s|)$ (Andrews, 1991; Newey and West, 1994). Note however that some of the estimated probabilities may be negative in finite samples although these probabilities are asymptotically positive. Antoine *et al.* (2007) propose a shrinkage procedure defined as a weighted average of the standard 2S-GMM's implied probabilities $(1/T)$ and the computed implied probabilities in order to guarantee the non-negativity property in finite samples.

In that respect, as shown by Brown and Newey (1998), Smith (2004) and Antoine *et al.*, (2007), an efficient estimator for the expectation of a function $h(z_t)$ can be obtained using implied probabilities. The usual estimator of $Eh(z_t)$ is given by $\frac{1}{T}\sum_{t=1}^{T} h(z_t)$. Then an efficient estimator, which makes use of the information content in the moment conditions, can be achieved by replacing the unconstrained probabilities $(1/T)$ with the implied probabilities (Eq. 3). The corresponding efficient estimator is $\sum_{t=1}^{T} p_t^{GMM}(\theta)h(z_t)$. As a result, a variance reduction is achieved by removing the correlation between the function $h(z_t)$ and the moment conditions $g(z_t,\theta)$ through the estimation of the expectation of $h(z_t)$ at the constrained implied probabilities $p_t^{GMM}(\theta)$:

$$\sum_{t=1}^{T} p_t^{GMM}(\theta)h(z_t) = \frac{1}{T}\sum_{t=1}^{T} h(z_t) - \frac{1}{T-p}\sum_{t=1}^{T} h(z_t)\left[J_{tT}(\theta) - \overline{J}_T(\theta)\right]' \hat{\Omega}_T(\theta)^{-1}\overline{g}_T(\theta)$$

where $\frac{1}{T-p}\sum_{t=1}^{T} h(z_t)\left[J_{tT}(\theta) - \overline{J}_T(\theta)\right]$ is a consistent estimator of the covariance matrix between $h(z_t)$ and the moment conditions $g(z_t,\theta)$ for standard kernels $\kappa(|t-s|)$.[3]

As a final remark, it is worth noticing that the CUE, defined by[4]

$$\hat{\theta}_T^{CUE} = \arg\min_{\theta\in\Theta} \frac{1}{T}\sum_{t=1}^{T} g(z_t,\theta)'\hat{\Omega}_T(\theta)^{-1}\frac{1}{T}\sum_{t=1}^{T} g(z_t,\theta),$$

uses the relevant constrained estimator of the Jacobian matrix by taking into account implied probabilities (Newey and Smith, 2004; Antoine *et al.*, 2007)—the Jacobian matrix is replaced by the residuals of its regression on the moment conditions.[5]

[3] In an i.i.d. setting, this estimator is semi-parametrically efficient (Chamberlain, 1987).

[4] The objective function is simultaneously minimized over θ and $\hat{\Omega}(\theta)$. In other words, the empirical variance-covariance matrix of the moment conditions replaces the fixed metrics of the 2S-GMM objective function, in which a norm of empirical moments is minimized. For further details, see Hansen, Heaton and Yaron (1996).

[5] The CUE has important advantages over the conventional 2S-GMM estimator. First, unlike the 2S-GMM estimator, the CUE does not depend on the normalization of the moment conditions. Second, in contrast to

3 Bias-corrected 3-step estimators

In this section, we present two smoothed 3S-EEL estimators as well as their bias-corrected versions. Then we show that both 3-step (bias-corrected) smoothed estimators are asymptotically equivalent and share the same higher-order equivalence as the smoothed empirical likelihood estimator up to an order $\mathcal{O}_p(T^{-3/2})$). Their bias-correction versions are asymptotically unbiased up to order T^{-1}.

The 3S-EEL estimator, proposed by Antoine *et al.* (2007) in the i.i.d. context, has the two interesting properties of being efficient with minimal asymptotic higher-order bias, like the EL estimator, and of preserving the user-friendly features of least squares. Unlike the standard 2S-GMM estimator, it uses all information contained in the moments conditions to estimate θ and thus improve the estimation of the optimal selection of estimating equations. Generally speaking, these equations correspond to the FOC of the EL estimator in the i.i.d. context (Newey and Smith, 2004) given some efficient estimators of the Jacobian and the optimal weighting matrices. Let $\hat{\theta}_T$ be an efficient GMM estimator, say the 2S-GMM estimator, the 3S-EEL estimator is defined as the solution of the following p equations

$$\left[\widetilde{G}_T(\hat{\theta}_T)\right]' \left[\widetilde{\Omega}_T(\hat{\theta}_T)\right]^{-1} \frac{1}{T}\sum_{t=1}^{T} g(z_t, \hat{\theta}_T^{3S}) = 0 \tag{4}$$

where $\widetilde{G}_T(\hat{\theta}_T)$ and $\widetilde{\Omega}(\hat{\theta}_T)$ are efficient estimators of the Jacobian and the variance-covariance matrices of the moment conditions

$$\widetilde{G}_T(\hat{\theta}_T) = G_T(\hat{\theta}_T) - Cov_T\left[\frac{\partial g}{\partial \theta'}(z_t, \hat{\theta}_T), g(z_t, \hat{\theta}_T)\right] \hat{\Omega}_T(\hat{\theta}_T)^{-1}\overline{g}_T(\hat{\theta}_T)$$

and

$$\widetilde{\Omega}_T(\hat{\theta}_T) = \hat{\Omega}_T(\hat{\theta}_T) - Cov_T\left[g(z_t, \hat{\theta}_T)g(z_t, \hat{\theta}_T)', g(z_t, \hat{\theta}_T)\right] \hat{\Omega}_T(\hat{\theta}_T)^{-1}\overline{g}_T(\hat{\theta}_T)$$

with $G_T(\hat{\theta}_T) = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial g}{\partial \theta'}(z_t, \hat{\theta}_T)$, $\hat{\Omega}_T$ is a standard consistent estimator of the covariance matrix of the moment conditions, and $Cov_T(X, Y)$ is a consistent estimator of the covariance matrix between the matrices (or vectors) X and Y.

More precisely, these efficient estimators of the Jacobian and the covariance matrices of the moment conditions $g(z_t, \theta)$ are equivalent to

$$\widetilde{G}_T(\hat{\theta}_T) = \sum_{t=1}^{T} p_t^{CUE}(\hat{\theta}_T)\frac{\partial g}{\partial \theta'}(z_t, \hat{\theta}_T), \tag{5}$$

the 2S-GMM estimator, Newey and Smith (2004) in the i.i.d. case and Anatolyev (2005) with dependent data show that the higher-order asymptotic bias of the CUE does not increase with the number of over-identifying restrictions. In addition, Newey and Smith (2004) and Anatolyev (2005) demonstrate that the CUE has the same minimal higher-order bias as the EL estimator if the third moments of the moment conditions are null. At the same time, it is more sensitive to initial conditions than the 2S-GMM estimator and it uses an unconstrained estimator of the weighting matrix.

and

$$\widetilde{\Omega}_T(\hat{\theta}_T) = \sum_{t=1}^{T} p_t^{CUE}(\hat{\theta}_T) g(z_t, \hat{\theta}_T) g(z_t, \hat{\theta}_T)' \tag{6}$$

with

$$p_t^{CUE}(\theta) = \frac{1}{T} - \frac{1}{T-p} \left[g(z_t, \theta) - \overline{g}_T(\theta) \right]' \hat{\Omega}_T(\theta)^{-1} \overline{g}_T(\theta) \tag{7}$$

where the p_t^{CUE}'s are the solution of the minimization problem between the constrained empirical distribution and the unconstrained empirical distribution $(1/T)$ according to the Chi-square distance—these implied probabilities being those of the CUE (see Antoine *et al.*, 2007).

To derive our proposed estimators in the time-series context, we need to redefine Eq. (4), (5), (6), and (7). Indeed moment conditions have to be smoothed with an appropriate kernel to account for the presence of temporal dependence. Following Smith (2004), the smoothed moment conditions are defined by

$$g_{tT}(\theta) = \frac{1}{S_T} \sum_{s=t-T}^{t-1} k\left(\frac{s}{S_T}\right) g(z_{t-s}, \theta)$$

where $t = 1, \ldots, T$, S_T is a bandwidth parameter with $S \to \infty$ as $T \to \infty$, and $k(\cdot)$ is a kernel function with $k_j = \int_{-\infty}^{\infty} k(a)^j da$.[6] Using the uniform kernel proposed by Kitamura and Stutzer (1997), one has

$$g_{tT}(\theta) = \frac{1}{2K_T + 1} \sum_{s=-K_T}^{K_T} g(z_{t-s}, \theta),$$

and $k_1 = k_2 = 1$.

On the other hand, the smoothed derivatives of the moment conditions are given by

$$G_{tT}(\theta) = \frac{1}{S_T} \sum_{s=t-T}^{t-1} k\left(\frac{s}{S_T}\right) \frac{\partial g}{\partial \theta'}(z_{t-s}, \theta).$$

Therefore our smoothed 3-step estimators proceed in the same way as the EL estimator, but for its smoothed version—the Smoothed Empirical Likelihood (SEL) estimator. More specifically, the FOC of the SEL estimator are shown to imply the following p equations (Smith, 2004; Anatolyev, 2005)

$$\left[\sum_{t=1}^{T} p_t^{SEL}(\hat{\theta}_T^{SEL}) G_{tT}(\hat{\theta}_T^{SEL}) \right]' \left[S_T \sum_{t=1}^{T} p_t^{SEL}(\hat{\theta}_T^{SEL}) g_{tT}(\hat{\theta}_T^{SEL}) g_{tT}(\hat{\theta}_T^{SEL})' \right]^{-1} \frac{1}{T} \sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{SEL}) = 0. \tag{8}$$

[6] Smith (2004) presents examples of appropriate kernels to smooth the moment conditions and their resulting induced kernel for the estimation of the variance-covariance matrix of the moment conditions. For instance, the uniform kernel proposed by Kitamura and Stutzer (1997) induces the Bartlett kernel for the estimation of the variance-covariance matrix.

with $p_t^{SEL}(\hat{\theta}_T^{SEL}) = 1/\left(1 - \hat{\lambda}' g_{tT}(\hat{\theta}_T^{SEL})\right)$ where $\hat{\lambda}$ is the vector of the Lagrange multiplier associated to the respective moment conditions (see Smith 2004). The evaluation of the Jacobian at the implied probabilities $p_t^{SEL}(\hat{\theta}_T^{SEL})$ allows to remove a component appearing in the asymptotic bias formula of order T^{-1} for the 2S-GMM. This bias component originates from the correlation between the Jacobian and the moment conditions. Newey and Smith (2004) show for models estimated by instrumental variables in i.i.d. context that this bias component grows linearly with the number of overidentifying restrictions. Note that this bias component is also absent for the CUE. Moreover Anatolyev (2005) establishes that the evaluation of the weighting matrix at the implied probabilities using an appropriate kernel removes the bias component involved by the third moments of the moment conditions. This bias component appears in the asymptotic bias formula of order T^{-1} for the 2S-GMM and the CUE.

Finally, implied probabilities have to be defined for weakly dependent data. Following Section 2, we draw from the property that implied probabilities have a closed form in an Euclidian space for the CUE (Antoine *et al.*, 2007). Therefore, using the smoothed moment conditions, it is straightforward to obtain the following definition[7]

Definition 1 *The implied probabilities corresponding to the smoothed CUE (SCUE) are given by*

$$p_t^{SCUE}(\theta) = \frac{1}{T} - \frac{1}{T-p}\frac{S_T}{k_2}\left[g_{tT}(\theta) - \overline{g}_T(\theta)\right]' \hat{\Omega}_T(\theta)^{-1}\overline{g}_T(\theta) \tag{9}$$

where $\overline{g}_T(\theta) = \frac{1}{T}\sum_{t=1}^T g_{tT}(\theta)$ *and* $\hat{\Omega}_T(\theta) = \frac{1}{T}\frac{S_T}{k_2}\sum_{t=1}^T g_{tT}(\theta)g_{tT}(\theta)'$ *is a consistent and positive definite estimator of the variance-covariance matrix* Ω*, and* k_2 *can be replaced by its empirical counterpart (see Smith, 2004).*

Consider again the estimation of the expectation for a scalar function $h(z_t)$ as in Section 2 but in the time series context. The smoothed version of the scalar function, $h(z_t)$, is now given by

$$h_{tT}(z_t) = \frac{1}{S_T}\sum_{s=t-T}^{t-1} k\left(\frac{s}{S_T}\right) h(z_{t-s}),$$

and an efficient estimator, which uses the information content of the moment conditions $g(z_t, \theta)$, has the following expression

$$\sum_{t=1}^T p_t^{SCUE}(\theta) h_{tT}(z_t) = \frac{1}{T}\sum_{t=1}^T h_{tT}(z_t) - \frac{1}{T-p}\frac{S_T}{k_2}\sum_{t=1}^T h_{tT}(z_t)\left[g_{tT}(\theta) - \overline{g}_T(\theta)\right]' \hat{\Omega}_T(\theta)^{-1}\overline{g}_T(\theta)$$

where $\frac{1}{T-p}\frac{S_T}{k_2}\sum_{t=1}^T h(z_t)\left[g_{tT}(\theta) - \overline{g}_T(\theta)\right]'$ is a consistent estimator of the covariance matrix between $h(z_t)$ and the moment conditions $g(z_t, \theta)$ (Smith 2004, Theorem 3.1).

[7] These implied probabilities can be derived by using results in Smith (2004). In his notation, the smoothed CUE corresponds to the SGEL criteria for $\rho(v) = -(1+v)^2/2$ for $v = k\lambda' g_{tT}(\theta)$, $k = k_1/k_2$ and λ is a vector of auxiliary parameters. The corresponding implied probabilities are then given by the expression (3.1) in Smith (2004).

Given Definition 1, we can now present the 3-step estimators based on the FOC of the SEL (Eq. 8). The first smoothed 3-step estimator is the one proposed by Antoine *et al.* (2007) but for weakly dependent data. As stated in Definition 2, the 3S-EEL estimator solves the p equations (8) after evaluating the Jacobian and the weighting matrices at an efficient estimator of θ, say the 2S-GMM estimator.

Definition 2 *The smoothed 3S-EEL estimator, $\hat{\theta}_T^{3S}$, is the solution of the following p equations:*

$$\left[\sum_{t=1}^{T} p_t^{SCUE}(\hat{\theta}_T) G_{tT}(\hat{\theta}_T)\right]' \left[S_T \sum_{t=1}^{T} p_t^{SCUE}(\hat{\theta}_T) g_{tT}(\hat{\theta}_T) g_{tT}(\hat{\theta}_T)'\right]^{-1} \frac{1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{3S}) = 0 \quad (10)$$

where $\hat{\theta}_T$ is an efficient estimator of θ and $p_t^{SCUE}(\cdot)$ is defined in eq. (9).

In contrast to the SEL estimator, the implied probabilities corresponding to the Chi-square metric (evaluated at an efficient estimator of θ) are used to estimate the Jacobian and the weighting matrices. The resulting estimator can then be computed much more easily than other smoothed GEL estimators. The solution is quite straightforward whether the smoothed moment conditions, $\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{3S})$, are either linear or nonlinear. At the same time, the use of an efficient Jacobian matrix, which is evaluated at an efficient estimator resulting from a preceding estimation step, might only lead to correct partially the finite sample bias component arising from the correlation between the Jacobian and the moment conditions. It turns out that this conjecture is confirmed by our simulation experiments below. In that respect, we propose an alternative 3-step estimator, denoted $\hat{\theta}_T^{3SW}$, where the Jacobian is left unrestricted and the weighting matrix is computed with implied probabilities evaluated at an efficient estimator resulting from a preceding estimation step (usually the 2S-GMM). Let us now define the smoothed 3SW-EEL estimator.

Definition 3 *The smoothed 3SW-EEL estimator, $\hat{\theta}_T^{3SW}$, is the solution of the following p equations:*

$$\left[\sum_{t=1}^{T} p_t^{SCUE}(\theta_T^{3SW}) G_{tT}(\hat{\theta}_T^{3SW})\right]' \left[S_T \sum_{t=1}^{T} p_t^{SCUE}(\hat{\theta}_T) g_{tT}(\hat{\theta}_T) g_{tT}(\hat{\theta}_T)'\right]^{-1} \frac{1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{3SW}) = 0. \quad (11)$$

To some extent, this estimator is more in the spirit of the traditional 2-step GMM—only the weighting matrix is evaluated at the estimator obtained at the preceding estimation step. While this estimator is computational more demanding than the one in Definition 2, it remains less demanding than the SEL estimator.

We now discuss the asymptotic properties of the estimator presented in Definitions 2 and 3. Both estimators are asymptotically higher-order equivalent to the SEL estimator up to order $\mathcal{O}_p(T^{-3/2})$. Indeed, starting from Anatolyev (2005), we show that the second-order asymptotic bias of our proposed estimators lacks some components with respect to the 2S-GMM estimator. More specifically, both estimators remove the bias component resulting from the correlation between the moment conditions and their derivatives and also remove the bias component associated with third moments by using an appropriate choice of the kernel.[8] Finally, even with moment conditions serially uncorrelated but not i.i.d. across time, Anatolyev (2005) shows that the SEL tends to reduce the bias—a property shared by our smoothed 3-step EEL estimators.

In that respect, the next proposition sets forth the higher-order efficiency equivalence between the SEL estimator and the 3SW-EEL estimator.

Proposition 1 *Under the Assumptions A1 to A7 in the Appendix, the smoothed three-step estimator, $\hat{\theta}_T^{3SW}$, defined as the solution of the p equations*

$$\hat{\theta}_T^{3SW} - \hat{\theta}_T^{SEL} = \mathcal{O}_p(T^{-3/2})$$

and thus achieves the same higher-order efficiency as the smoothed empirical likelihood estimator.

Proof: see Appendix.

This higher-order efficiency also holds for the time series extension of the 3S-EEL. The characterization of the asymptotic higher-order properties of the smoothed 3-step estimators in Proposition 1 leads to several remarks. In principle, these estimators can also be computed with the unsmoothed moment conditions. In this case, the corresponding implied probabilities are defined in Eq. (3)—they insure the consistency of the Jacobian and the variance-covariance matrices estimators of the moment conditions. Second, the smoothed 3-step estimators share the same higher-order asymptotic properties as the SEL estimator for certain class of kernels, as for instance the uniform kernel proposed by Kitamura and Stutzer (1997). In the sequel, our simulation experiments and the application are based on the uniform kernel. From a practical view, the smoothing parameter K_T is chosen according to the data-dependent procedure proposed by Newey and West (1994).[9] Third, the smoothed 3S-EEL, 3SW-EEL and the SEL estimators have the same bias-order, namely $\mathcal{O}(T^{-1})$, so that the higher-order asymptotic derivations in Anatolyev (2005) allow us for proposing a bias-corrected version of these estimators. The next proposition gives the corresponding expression for the smoothed 3SW-EEL estimator (Definition 3). The same result applies for the smoothed 3S-EEL estimator (Definition 2).

[8] The smoother proposed by Kitamura and Stutzer (1997) belongs to this class of appropriate kernels.

[9] More precisely, K_T is chosen by using the relationship established by Smith (2004) between the truncated kernel for the moment conditions and the corresponding induced Bartlett kernel. K_T is thus fixed to the integer value of $(m_T - 1)/2$ where m_T is the lag length chosen by the data-driven procedure of Newey and West (1994).

Proposition 2 *Under Assumptions A1 to A7 in the Appendix, a consistent estimator of the asymptotic bias of order* T^{-1} *is given by:*

$$\widehat{Bias}(\hat{\theta}_T^{3SW}) = \hat{B}_{G\Xi g}/T + \hat{B}_{\partial^2 g}/T$$

where $\hat{B}_{G\Xi g}$ *and* $\hat{B}_{\partial^2 g}$ *are consistent estimators of:*

$$B_{G\Xi g} = \Xi \sum_{u=-\infty}^{\infty} E\left[\frac{\partial g}{\partial \theta'}(z_t, \theta)\Xi g(z_{t-u}, \theta)\right]$$

$$B_{\partial^2 g} = \Xi \sum_{j=1}^{p} E\left[\frac{\partial^2 g}{\partial \theta' \partial \theta_j}(z_t, \theta)\frac{\Sigma}{2}e_j\right]$$

and e_j *is the* j*th column of the identity matrix of order* p, $\Sigma = \left(G'\Omega^{-1}G\right)^{-1}$, $\Xi = \Sigma G'\Omega^{-1}$, $G = E\left[\frac{\partial g}{\partial \theta'}(z_t, \theta)\right]$ *and* $\Omega = \sum_{s=-\infty}^{\infty} E\left[g(z_t, \theta)g(z_{t-s}, \theta)'\right]$. *The bias corrected smoothed three-step estimators* $\hat{\theta}_T^{3SWc}$ *defined as* $\hat{\theta}_T^{3SWc} = \hat{\theta}_T^{3SW} - \widehat{Bias}(\hat{\theta}_T^{3SW})$ *are asymptotically unbiased up to order* T^{-1}.

Proof: see Appendix.

These two terms represent the asymptotic bias for a GMM estimator based on the infeasible optimal combination of moment conditions. Consistent estimators of $B_{G\Xi g}$ and $B_{\partial^2 g}$ are obtained following an appropriate replacement of moment conditions or their derivatives by their respective smoothed versions (see Lemmas 2 and 3 in Anatolyev, 2005). Thus, $\widetilde{G}_T = \sum_{t=1}^{T} p_t^{SCUE}(\hat{\theta}_T^{3SW})G_{tT}(\hat{\theta}_T^{3SW})$, $\widetilde{\Omega}_T = \frac{S_T}{k_2}\sum_{t=1}^{T} p_t^{SCUE}(\hat{\theta}_T^{3SW})g_{tT}(\hat{\theta}_T^{3SW})g_{tT}(\hat{\theta}_T^{3SW})'$ and a consistent estimator of $\sum_{u=-\infty}^{\infty} E\left[\frac{\partial g}{\partial \theta'}(z_t, \theta)\Xi g(z_{t-u}, \theta)\right]$ is given by:

$$\frac{S_T}{k_2}\sum_{t=1}^{T} p_t^{SCUE}(\hat{\theta}_T^{3SW})G_{tT}(\hat{\theta}_T^{3SW})\widetilde{\Xi}_T g_{tT}(\hat{\theta}_T^{3SW})$$

where $\widetilde{\Xi}_T = \left(\widetilde{G}_T'\widetilde{\Omega}_T^{-1}\widetilde{G}_T\right)^{-1}\widetilde{G}_T'\widetilde{\Omega}_T^{-1}$ (see Lemma 3b in Anatolyev (2005)). Finally, a consistent estimator of the second bias term is obtained with $\widetilde{\Sigma}_T = \left(\widetilde{G}_T'\widetilde{\Omega}_T^{-1}\widetilde{G}_T\right)^{-1}$ and the second partial derivative of the smoothed moment conditions with respect to the parameter vector θ. It is worth noticing that the bias terms can also be estimated at the usual weigth $1/T$ instead of the constrained implied probabilities. However such estimators are not efficient compared with the ones proposed here. We conjecture that efficient estimators of the bias terms would probably improve the small sample performances of each bias-corrected estimator.

4 Simulation experiments

In this section, we examine the finite sample properties of the CUE, the 2S-GMM, and the smoothed (bias-corrected) 3S-EEL and 3SW-EEL estimators.

4.1 The data generating process

We assume that the data generating process (DGP) is based on the (hybrid) quasi-structural form of a univariate rational expectations model, as for instance any log-linearized Euler equation in a dynamic stochastic stochastic general equilibrium model. Following Mavroiedis (2004, 2005), and Nason and Smith (2005), the forcing variable is driven by an autoregressive process of order p. The dynamic specification is thus given by

$$y_t = \gamma_f E_t y_{t+1} + \gamma_b y_{t-1} + \lambda x_t + \epsilon_t$$
$$\rho(L) x_t = v_t$$

where $\rho(L) = \rho_1 + \rho_2 L + \cdots \rho_p L^p$, γ_f, γ_b and λ are generally nonlinear functions of some structural (or deep) parameters, say $\theta \in \Theta$, ϵ_t is an exogenous shock with zero mean and variance σ_ϵ, and v_t is the innovation process. The variance-covariance matrix of the error terms is defined by

$$\Sigma = \begin{pmatrix} \sigma_\epsilon^2 & \sigma_{\epsilon v} \\ \sigma_{\epsilon v} & \sigma_v^2 \end{pmatrix}.$$

The estimation methods use the sample version of the following moment conditions

$$E\left[Z_t \left(y_t - \lambda x_t - \gamma_f y_{t+1} - \gamma_b y_{t-1}\right)\right] = 0 \quad (12)$$

where the vector Z_t denotes the set of appropriate instruments.

As well-explained by Nason and Smith (2005), identification requires predictability of future forcing variable values beyond that provided by the current ones, or current or lagged endogenous variable. Since x_t follow a p-order autoregressive process, $p \geq 2$ is necessary for identification and $p \geq 3$ for over-identification.[10] Consequently, for an AR(1) process, the parameters of our DGP cannot be identified by GMM. In the sequel, we assume that the forcing variable is driven by an AR(2) process and the reduced-form is thus

$$\begin{aligned} y_t &= \delta_1 y_{t-1} + \alpha_0 x_t + \alpha_1 x_{t-1} + \alpha_\epsilon \epsilon_t \\ x_t &= \rho_1 x_{t-1} + \rho_2 x_{t-2} + v_t \end{aligned}$$

where $\delta_1 = \frac{1-\sqrt{1-4\gamma_b\gamma_f}}{2\gamma_f}$, $\alpha_0 = \frac{\lambda}{\Delta\delta_2\gamma_f}$, $\alpha_1 = \alpha_0\frac{\rho_2}{\delta_2}$, $\alpha_\epsilon = \frac{1}{\delta_2\gamma_f}$, and δ_2 and Δ are respectively given by $\frac{1+\sqrt{1-4\gamma_b\gamma_f}}{2\gamma_f}$ and $1 - \frac{\rho_1}{\delta_2} - \frac{\rho_1}{\delta_2^2}$.

However, even though the necessary condition of identification is respected, the strength of

[10] For an extensive discussion, see Mavroiedis (2004, 2005) and Nason and Smith (2005).

identification should be studied more precisely.[11] In particular, the model is under-identified when $\lambda = 0$—the t-statistics for the hypothesis $H_0 : \lambda = 0$ does not have an asymptotically normal distribution under the null (Dufour, 1997, 2003). In that respect, we follow the approach developed by Mavroiedis (2004, 2005). Indeed Mavroieidis (2004, 2005) examines the weak identification issue by determining the concentration parameter of the reduced form, which depends both on the quasi-structural parameters λ, γ_f, γ_b, and σ_ϵ^2, and on the nuisance parameters ρ_i, σ_v^2, and $\sigma_{v\epsilon}$. Moreover, it is argued that the concentration parameter is invariant to re-scaling of the data so it depends only on $\sigma_v^2/\sigma_\epsilon^2$ and $\sigma_{v\epsilon}$. In particular, the strength of identification is increasing in $\sigma_v^2/\sigma_\epsilon^2$. Therefore, to shed some light on the weak identification issue, we report in our Monte-Carlo experiments the value of the concentration parameter.[12]

4.2 Finite and large sample properties of the estimators

We report Monte Carlo evidence on the quasi-structural parameters, λ, γ_f, and γ_b. According to the theoretical model, these parameters satisfy the restrictions γ_f, $\gamma_b \geq 0$, $\gamma_f + \gamma_b < 1$ and $\lambda \geq 0$ (see Buiter and Jewitt, 1989, and Galí and Gertler, 1999).[13] These restrictions imply that the reduced-form is determinate, and, thus, the backward- and forward-looking parameters are only partially identified when $\lambda = 0$. Three sets of parameters are of particular interest. On the one hand, according to our empirical application in Section 5, our benchmark parameters are the ones estimated by Gali and Gertler (1999, Table 2) in the case of the NKPC. Hence, $\gamma_f = .591$,$\gamma_b = .378$, $\lambda = .015$. On the other hand, we assume that the DGP of y_t is mostly forward-looking (respectively backward-looking), e.g. $\gamma_f = .850$,$\gamma_b = .100$, $\lambda = .015$ (respectively $\gamma_f = .100$,$\gamma_b = .850$, $\lambda = .015$).[14] For each parameter set, two cases are worth studying. The first case (Case I) assumes that the model is well-identified. The AR(2) parameters for x_t are set to $\rho_1 = .9(1-\rho_2)$ and $\rho_2 = -.65$. with $\sigma_v^2/\sigma_\epsilon^2 = 8$. The error terms ϵ_t and v_t are drawn from a bivariate normal distribution with standard deviations $\sigma_\epsilon = .05$ and $\sigma_v = .4$. The correlation coefficient between the error terms takes respectively the values 0.5, 0, and -0.5. In the second case (Case II), the parameters of the AR(2) are determined such that all DGPs are weakly identified. Those values are $\rho_1 = .9(1 - \rho_2)$ and $\rho_2 = -.65/\sqrt{(}T)$ where T denotes the sample size.

[11]It is worth noticing that the mapping from the structural parameters θ to $(\lambda, \gamma_f, \gamma_b)$ is generally not invertible, i.e. the structural parameters are not globally identified (Ma, 2002). In addition, there are regions in the admissible space in which these structural parameters become locally unidentified. For instance, see Dufour, Khalaf and Kichian (2005).

[12]The derivation of the concentration parameter is provided in the technical report available upon request.

[13]We also consider the case in which $\gamma_f + \gamma_b = 1$. Following Blanchard and Kahn (1980), two situations can be encountered. When $\gamma_f \leq 0.5$, the solution of the characteristic polynomial is unique, but y_t is a non-stationary process regardless the dynamics of x_t. When $\gamma_f > 0.5$ and second-order stationary conditions on the forcing variable hold, the existence of a stationary solution is guaranteed, but there are in fact infinitely many solutions characterized by sunspot shocks. Results are not reported here but are available upon request.

[14]It turns out that the last set of parameters leads to a concentration nearby zero so that they are weakly identified.

To investigate the effects of the number of instruments regarding the small sample performances of the estimators, we consider instrument sets including $K/2$ lags values of y_t and x_t where K, the number of instruments, equals respectively $8, 16$ and 24.[15] The sample size is 160, e.g. the sample size of our empirical application in Section 5. All results reported below are based on 5,000 simulation repetitions. For each repetition, we calculate the CUE, the 2S-GMM, the smoothed 3S-EEL, the smoothed (bias-corrected) 3SW-EEL (3SWc-EEL) estimators of the quasi-reduced form parameters (γ_b, γ_f, λ).[16] We then calculate the mean bias and the root mean squared error (RMSE) of the estimators over the 5,000 samples.[17]

From a computational view, we use the numerical optimization routine *fminsearch.m*, which is a part of the "Optimization toolbox" in Matlab. We discard cases where the routine failed to converge.[18] Initial values were set to the true ones. While the smoothed (bias-corrected) 3S-EEL and the 2S-GMM estimators were immune to such an initialization, the CUE often fails to converge or yields large implausible values of the parameters. This numerical instability of GEL-based estimators is well-known in the literature and has been documented among others by Guggenberger and Hahn (2005), Anderson and Kunitomo (2005). Consequently, the CUE may display higher mean bias and RMSE, especially in finite samples. Note finally that the variance-covariance matrix of the moment conditions is estimated using the automatic lag procedure of Newey and West (1994). In unreported results, we also test the sensitivity of our results with respect to the number of lags used in the computation of the variance-covariance matrix of the moment conditions. In particular, we consider a fixed window up to twelve lags and the procedure of West (1997). In the latter, if the model is correctly specified, the error term of y_t follows an MA(1) process.[19] Overall, results only marginally improve.

We now discuss the relative performance of each estimator under strong identification (case I) and weak identification (case II). We also discuss the effect of T and K on the performance of the estimators. Table 1 reports the small sample simulation results (T=160) for the first set of parameter vector. Several points are worth discussing. First, the concentration parameter, denoted *cp*, clearly shows that the parameters are well-identified irrespective of the correlation coefficient, ρ, and the number of instruments, K.[20] Second, the bias of all estimators increases

[15] Results for $K = 4$ and 12 are available upon request.

[16] Results are not reported for the smoothed 3Sc-EEL estimator since the bias correction only changes marginally the results of the smoothed 3S-EEL estimator.

[17] We also calculate the median bias and the median absolute deviation. Results are not reported here but are available upon request.

[18] A fine grid search approach is recommended by Hansen et al. (1996) and Guggenberger (2006) for the CUE and the GEL estimator to circumvent the problem of numerical instability (or unreliability of minimization routines) often encountered with these estimators even in the case of a scalar parameter.

[19] However, the West procedure cannot be implemented for our smoothed 3S-EEL estimators since this requires to define the correct kernel in the smoothed moment conditions. We leave this issue for further research.

[20] The parameters are considered well identified for a minimum eigenvalue (concentration parameter) superior to 10.

with the number of instruments, with a more sizeable effect for the 2S-GMM estimator. This confirms that the Jacobian evaluated at the implied probabilities contributes to partially correct the effect of the number of overidentifying restrictions on the bias. This correction is even more pronounced for the 3SW-EEL and 3SWc-EEL estimators and thus provides some support of our conjecture in Section 3 (except for the backward-looking parameter in some cases)—leaving unrestricted the Jacobian matrix improves the statistical performances of 3-step estimators. Third, the smoothed 3-step estimators dominate the 2S-GMM estimator and the CUE in terms of mean bias and RMSE. In addition, the CUE has a larger RMSE than other estimators and performs rather poorly for the coefficient of the forcing variable. This heavy tails problem for the CUE is well-known in the literature and has been shown, among others, by Hansen, Heaton and Yaron (1996), Guggenberger and Hahn (2005), and Guggenberger (2005). Fourth, the relative magnitude of the mean bias (with respect to the true values) is far from being negligible for the forward-looking and the forcing variable parameters, except for, to some extent, when the number of instruments is small $K = 8$. This suggests that this small-sample bias may significantly distort standard GMM estimates of univariate (multivariate) rational expectations models in empirical applications. Finally, the correlation parameter ρ does not change the ordering of the estimators in terms of bias and RMSE.

[Insert Tables 1 and 2 around here]

Unsurprisingly, as the sample size increases (Table 2), $T = 500$, the mean bias and the RMSE significantly reduce for all estimators. Noticeably, as to be expected from theory, the CUE competes now very favorably in terms of mean bias with our proposed estimators—these being only asymptotically equivalent to the CUE. It is however at the expense of a higher RMSE, especially with respect to the 3SWc-EEL estimator. All in all, it turns out that our previous conclusions remain valid in large samples. We now assess the robustness of our simulation results using other parameters values.

Table 3 reports the small-sample simulation results when the DGP of y_t is mostly forward-looking, e.g. $\gamma_f = .850$,$\gamma_b = .100$, $\lambda = .015$. Three points are worth commenting. On the one hand, the mean bias and the RMSE are higher, except for the forcing variable parameter, than those reported in Tables 1 and 2 irrespective of the correlation parameter and the number of moment conditions. In other words, all estimators underestimate (overestimate) the forward-looking (backward-looking) coefficient. To some extent, this result can be understand by analyzing the reduced-form coefficients. The stable root of the characteristic polynomial of the endogenous variable y_t, δ_1, is less persistent (0.11) than the one in the first set case (0.57) whereas the coefficients of the current and the lagged forcing variable, α_0 and α_1, are roughly the same (0.09 (respectively -0.06) instead of 0.11 (respectively -0.07) for α_0 (respectively α_1). Consequently, *ceteris paribus*, the DGP is mainly driven by the statistical properties of the forcing variable and the correlation coefficient. Therefore, when $\rho = 0$ or 0.5, the CUE generally outperforms other estimators in terms of mean bias for γ_f, but still at the cost of a higher RMSE and poor finite sample statistical properties for the coefficient of the forcing variable. In contrast, when $\rho = -0.5$, the CUE is dominated by the 3-step estimators, especially the 3SW-EEL and the 3SWc-EEL estimators. At the same time, our proposed estimators often have better finite sample properties for the backward-looking coefficient and λ than the CUE and the 2S-GMM estimator. Finally, the smoothed 3SW-EEL estimator and its bias-corrected version generally overrule the smoothed 3S-EEL estimator in terms of mean bias and RMSE.

[Insert Tables 3 and 4 around here]

This interpretation remains valid when the sample size increases (Table 4). In particular, the large sample mean bias and RMSE of γ_b and γ_f is on average up to 2 times larger than those of our benchmark case, especially for the 2S-GMM estimator. In contrast, the bias significantly reduces for the coefficient of the forcing variable. Overall, the CUE performs better (in both criteria) than other estimators for the forward-looking coefficient. This conclusion remains valid for the backward-looking coefficient, except for the RMSE, and λ, when the number of instruments equal 8 or 16. Otherwise, our proposed estimators outperforms the 2S-GMM estimator and have an interesting bias-efficiency trade-off relative to the CUE.

Interestingly, the simulation results (Tables 5 and 6) differ when the DGP of y_t is mostly backward-looking, e.g. $\gamma_f = .100, \gamma_b = .850$, $\lambda = .015$. In contrast to previous Monte Carlo experiments, the concentration parameter is closer to zero, suggesting that parameters could be weakly identified. Moreover all estimators overestimate γ_f and underestimate λ and γ_b. Looking at the reduced-form coefficients, the stable root of the characteristic polynomial of the endogenous variable y_t, δ_1, is more persistent (0.94) than the one in our benchmark case (0.57) whereas the coefficients of the current and the lagged forcing variable, α_0 and α_1, are nearby zero (resp. 0.02 and -0.001). Therefore, *ceteris paribus*, the DGP assigns a small weight to the forcing variable and thus slightly depends on the correlation coefficient. In that respect, the CUE and the 3SW-EEL estimator have better finite and large sample properties than the 2S-GMM, the 3S-EEL, and the 3SWc-EEL estimators. While the CUE prevails asymptotically over the 3SW-EEL estimator (Table 6), the finite sample results show that the relative performance of the 3SW-EEL estimator with respect to the CUE depends on the parameter of interest and the number of instruments. Overall, the CUE has a large RMSE especially for the forcing variable λ.

[Insert Tables 5 and 6 around here]

We now turn to the weakly identified case (case II).[21] We report here the results when the DGP is mostly forward-looking. This allows us to describe the consequences of weak identification when we consider the most well-identified DGP in case I irrespective of the values of ρ and

[21] In contrast to Tables 5 and 6, weak identification arises here from the DGP of the forcing variable. In particular, this leads to weak instruments.

T and the necessary condition of identification stated in Section 4 is weakly respected. Since results are qualitatively similar, we only discuss the case when $\rho = 0$. Table 7 clearly shows that all estimators are significantly biased irrespective of the number of instruments (and the correlation parameter). Second, the RMSE also significantly augments for all estimators relative to the well-identified case. Moreover, the bias and the RMSE increase with the number of instruments since instruments are weak and thus do not convey reliable information. Finally, as to be expected from theory, the RMSE and the mean bias do not significantly fall with the sample size, i.e. the estimators do not converge to their true values. This result is illustrated in Table 8. This is in sharp contrast with the results reported in Table 3. We thus confirm the results of Stock and Wright (2000), Kleibergen (2002) and Mavroeidis (2004, 2005). Even if the comparison of the estimates is meaningless *per se* in the case of weak identification, it is worth noticing that our preferred estimators, the smoothed 3SW-EEL estimator and its bias corrected version, generally outperforms other estimators. Similar evidence is found in other cases.[22]

[Insert Tables 7 and 8 around here]

As a last experiment, we analyze the robustness of our results when the error term ϵ_t is drawn from a recentered $\chi^2(1)$. In that respect, we seek to evaluate the bias component arising from the third moments of the moment conditions. Table 9 presents the simulation results with a mostly forward-looking but well-identified DGP (Table 3). As to be expected from theory (Newey and Snith, 2004; Anatolyev, 2005), the bias and the RMSE grow in almost cases respective to the case with symmetric errors. However, this increase is less pronounced for the smoothed 3-step estimators than for the 2S-GMM estimator and the CUE.

[Insert Table 9 around here]

To sum up, our Monte Carlo simulations provide evidence that our proposed estimators are very competitive with respect to the 2S-GMM estimator and the CUE. In particular, results suggest that they perform better in terms of mean bias and RMSE than the 2S-GMM estimator. Second, among the proposed smoothed 3S-EEL estimators, the smoothed 3SW-EEL estimator and its bias-corrected version, generally have better finite and large sample properties than the time-series extension of the 3S-EEL estimator. In other words, leaving unrestricted the Jacobian matrix further improves the statistical performances of this class of estimators. Third, as T gets larger and larger, the CUE often performs very well However, as is well known, it is often at the expense of a larger RMSE, especially for γ_b and λ. Fourth, the finite sample bias encountered in univariate rational expectations models is far from being negligible even when the DGP is well-identified. As a result, this may significantly distort estimates and thus the corresponding interpretation (structural parameters).

[22] Results are available upon request.

5 Application

In this section, we report the results of the hybrid NKPC using the original dataset of Galí and Gertler (1960Q1-1997Q4).[23]

Since it burst onto the scene of mainstream monetary economics, the NKPC has been the focus of two important empirical debates. First, to what extent purely forward-looking pricing behavior can be reconciled with observed inflation persistence. Second, to what extent properly measured marginal costs affect inflation dynamics. Both issues are crucial for our ability to understand and predict movements in prices. They have recently been hotly debated, and for good reason. At the same time, some concerns have been raised on the methodology of Gali and Gertler (1999), and Gali *et al.* (2001, 2005). For instance, Rudd and Whelan (2005, 2006) or Lindé (2005) cast doubt on the validity of their GMM estimates. Dufour *et al.* (2005) and Mavroiedis (2004, 2005) stress the identification problem. Jondeau and LeBihan (2007) and Kurmann (2005) argue maximum likelihood estimator ought to be preferred. In that respect, we re-estimate the NKPC using our proposed estimators.

Briefly speaking, the NKPC advocated by Gali and Gertler (1999) and Gali, et al. (2001, 2005) is derived as follows. In a monopolistic environment, price setting decisions are driven by a modified version of the Calvo's (1983) staggering mechanism. In each period, each firm may have a fixed probability α not to adjust price. Among the firms facing a probability $1-\alpha$ to readjust prices, a proportion ω of firms does not optimally set their prices but do so in a pure backward-looking manner. The remaining fraction chooses their optimal price to maximize their expected discounted sum of profits.

The hybrid NKPC is then given by[24]

$$\pi_t = \lambda\kappa mc_t + \gamma_f E_t\pi_{t+1} + \gamma_b\pi_{t-1} + \epsilon_t \tag{13}$$

where $\lambda = ((1-\omega)(1-\alpha)(1-\alpha\beta))\,\phi^{-1}$, $\gamma_f = \beta\alpha\phi^{-1}$, $\gamma_b = \omega\phi^{-1}$, $\phi = \alpha + \omega\,[1-\alpha(1-\beta)]$, $\kappa = \frac{1}{1-\eta\mu}$, and $E_t\pi_{t+1}$ is expected inflation at time t, mc_t represents real marginal costs, β is the common subjective discount factor, μ is the firm's demand elasticities, and η is the elasticity of marginal cost. The constant, κ, represents a correction term of the forcing variable cost—the real marginal cost is a function of the average real marginal cost across firms.[25] All variables are expressed as a percentage deviation with respect to its steady state value.

The corresponding moment conditions are

$$E\left[Z_t\left(\pi_t - \gamma_f\pi_{t+1} - \gamma_b\pi_t - \lambda mc_t\right)\right] = 0$$

[23] See Data Appendix.

[24] The purely forward looking NKPC is nested in this specification.

[25] GGLS report estimates with $\kappa = 0.12$ and $\kappa = 1$ (see Section 4 in GGLS). In the sequel, we use $\kappa = 1$. The main conclusions relative to the estimation results do not change significantly with $\kappa = .12$

where Z_t is a vector of instruments dated t and earlier. We use two sets of instruments. The first one (GG) includes four lags each of inflation, the labor income share, the output gap, the long-short interest rate spread, wage inflation, and commodity price inflation, and corresponds to the one used by Galí and Gertler (1999). The second set (GGLS) reduces to four lags of inflation and two lags of the labor income share, output gap and wage inflation (Galí, Gertler and López-Salido, 2001).[26] In both cases, we use an automatic lag selection procedure as in Newey and West (1994).[27]

Before presenting our results, two points are worth noticing. First, given Definition 1, we perform a new specification test in addition to the usual J-statistic. Indeed, as suggested by Back and Brown (1993), implied probabilities may provide a useful diagnostic device. We thus define a statistic, which measures the discrepancy between the estimated probabilities and the unconstrained empirical probabilities $1/T$. The Implied Probability Statistic (IPS_T) is given by

$$IPS_T = \frac{k_2}{k_1^2 S_T} \sum_{t=1}^{T} \left(T p_t^{SCUE}(\hat{\theta}_T^{3S}) - 1 \right)^2$$

where $p_t^{SCUE}(\hat{\theta}_T^{3S})$'s are the implied probabilities defined in equation (9) but evaluated at the (bias-corrected) smoothed 3S-EEL or 3SW-EEL estimator. Under usual regularity conditions, in i.i.d. settings (S_T=1), this statistic is asymptotically distributed as $\chi^2(q-p)$ (see Theorem 1 in Baggerly, 1998, and Ramalho and Smith, 2005). In the Appendix, we show in the time-series context that this statistic is numerically equivalent to a J-statistic but computed with smoothed moment conditions and a centered weighting matrix (see Smith 2005). The IPS_T-statistic is then asymptotically first-order equivalent to the J-statistic computed with the standard 2S-GMM. Nevertheless, they can differ in small samples.

On the other hand, we reconsider the problem of identification in view of our simulation results. Obviously, we are far from the first ones to take interest in this problem. Especially relevant contributions on this issue include Ma (2002), Mavroiedis (2004, 2005, and 2007), Nason and Smith (2005), and Dufour, Khalaf and Kichian (2005). In that respect, we show in a companion paper (Guay and Pelgrin, 2007) that robust identification test statistics can be derived for the 2S-GMM and our proposed smoothed 3-step estimators in order to test a simple hypothesis on all parameters or a subvector of the parameters.[28]. The first test-

[26]In contrast to Galí and Gertler (1999), and Galí, Gertler, and López-Salido (2001), we use a real-time output-gap measure (quadratic trend) instead of a gap detrended using the full sample.

[27]As explained before, the error term is an MA(1) process (the one-step-ahead nature of the expected inflation forecasts), if the model is correctly specified. In particular, as Mavroiedis (2004) discussed, the presence of higher-order autocorrelation in the error term suggests that (i) the instruments used for estimation are invalid or (ii) the model's dynamic structure is misspecified, which adversely alters GMM estimation and inference. In that respect, we implement the higher-order autocorrelation test of Cumby and Huizinga (1992) and find little support to excess serial correlation.

[28]See the technical report

statistic is a straightforward extension of the results in Kleibergen (2005). Indeed, we make use of the results of Back and Brown (1993) and replace the unconstrained Jacobian matrix by the Jacobian matrix evaluated at the implied probabilities. The test statistic is then given as a quadratic form in the resulting score vector evaluated at the hypothesized parameter vector, say θ, and re-normalized at the appropriate rate. As a result, the test-statistic is asymptotically χ^2 distributed. We proceed in the same way to derive the robust identification statistics for our smoothed 3-step estimators. The equivalence between the LM-statistic defined in Guggenberger and Smith (2007) and ours is then straightforward to show. Since both test-statistics are asymptotically pivotal, the level of these tests should not vary too much in small samples under weak identification. In the sequel, we use these test-statistics for the null $H_0 : \lambda = 0$.

We now discuss our empirical results. We only report estimates of the reduced-form parameters (Tables 10 and 11). Several points are worth commenting. First, parameter estimates in Tables 10 and 11 are broadly in line with the results reported by Galí and Gertler (1999), and Galí, Gertler, and López-Salido (2001).[29] Using the large instrument set (GG), results provide significant support that the real marginal cost is the relevant forcing variable for the dynamics of inflation. In contrast, the use of the small instrument set (GGLS) generally leads to the statistical irrelevance of the forcing variable at the conventional 5% nominal size.[30] Second, estimates are quite similar among the proposed methods in Tables 10 and 11. However, as the number of instruments further reduces, we find differences across the estimators. Especially, the CUE and the smoothed 3-step estimators favor a more forward-looking dynamics than the 2S-GMM estimator.[31] Third, we cannot rule out weak-identification of the forcing variable coefficient irrespective of the number of instruments. Indeed the p-values of the K-based statistics provide strong evidence that we cannot reject the null, $H_0 : \lambda = 0$, at the 5% or 10% level. In that respect, our results support those of Mavroiedis (2007) and Dufour et al. (2006)—the forcing variable is weakly identified.[32] All in all, we find that the forward-looking component outweighs the backward dynamics and that weak identification is an issue for the forcing variable.

Finally there is evidence that the over-identifying restrictions are not rejected at conventional 5% level using either the J_T or the IPS_T statistics for the GGLS instrument set. However, these statistics differ greatly for the GG instrument set (e.g. with a large num-

[29] The estimation results are not exactly the same for two reasons. First, we use the Newey-West (1994) data driven procedure for the estimation of the optimal weighting matrix instead of fixing arbitrarily the number of lags to 12 as in Galí and Gertler (1999), and Galí, Gertler, and López-Salido (2001). Second, real-time output gap is used in the instrument set instead of a detrended gap over the full sample.

[30] In both cases, we re-estimate the NKPC using the restriction $\gamma_b + \gamma_f = 1$. Results only marginally differ and thus our conclusions are robust to this constraint. Results are available upon request.

[31] These results are consistent with those of Mavroeidis (2007, Table 1). Results are available upon request.

[32] We also test the null hypothesis $H_0 : \gamma_b = 0$ and $H_0 : \gamma_f = 0$. Evidence is mixed and depends on the set of instruments.

ber of instruments). On the one hand, the J_T statistic reported for all 3-step estimators is based on the smoothed moment conditions and an uncentered covariance matrix estimator. On the other hand, the IPS_T makes use of a centered covariance matrix estimator and is numerically equivalent to a "recentered" J_T statistic.[33] In that respect, the IPS_T statistic is in accordance with the contribution of Hall (2000), which advocates the use of a covariance estimator in mean deviation in order to increase the power of the overidentifying test. Hence the observed difference in Table 10 suggests a misspecification problem given that the conventional J_T statistic and the IPS_T statistic should be closed under the null that the moment conditions are not violated. To investigate more closely this issue, Figure 1 reports for both instrument sets the implied probabilities evaluated at the 3S-EEL estimator and their unconstrained counterparts.[34] For the GG instrument set, implied probabilities display large swings along the sample. In particular, we observe a substantial discrepancy between the constrained and unconstrained ($1/T$) implied probabilities in the seventies and the nineties. Moreover, several values are negative. As pointed out by Schennack (2007), negative values are more likely to occur in small sample and even asymptotically under misspecification. In contrast, with the small instrument set, implied probabilities deviate less from unconstrained probabilities and only few values are weakly negative. Therefore, misspecification is an issue for large instrument sets.

6 Conclusion

In this paper, we investigate the information content of implied probabilities (Back and Brown, 1993) to improve estimation in unconditional moment conditions models. We propose two smoothed (bias-corrected) 3S-EEL estimators for weakly dependent data. The first one is the time series extension of the 3S-EEL proposed by Antoine, Bonnal, and Renault (2007) in the i.i.d. context—it solves the FOC of the SEL estimator given the SCUE implied probabilities and some efficient estimators of the Jacobian and the optimal weighting matrices. In contrast, the second estimator only uses a weighting matrix computed with implied probabilities evaluated at an efficient estimator resulting from a preceding estimation step and the Jacobian is left unrestricted. Both estimators achieve a higher-order equivalence to the SEL (up to an order $\mathcal{O}_p(T^{-3/2})$)—their bias corrected versions are asymptotically unbiased up to order T^{-1}. Finally, these estimators avoid to solve a computational demanding saddle point problem as in the class of GEL estimators.

A Monte-Carlo study reveals that the finite sample properties of our new estimators are very competitive with respect to the 2S-GMM estimator and the CUE. Moreover, among the proposed smoothed 3-step estimators, the smoothed 3SW-EEL estimator and its bias-

[33] See the Appendix

[34] Implied probabilities computed with the other estimators (2S-GMM, CUE, 3Sc-EEL, 3SW-EEL and 3SWc-EEL) are similar.

corrected version generally have better finite and large sample properties than the time-series extension of the 3S-EEL estimator. In other words, leaving unrestricted the Jacobian matrix further improves the statistical performances of this class of estimators. As an application, we re-estimate the NKPC in the US and find evidence that the inflation dynamics is mostly forward-looking and driven by the real marginal cost. However, weak identification cannot be ruled out and misspecification is an issue for a large set of instruments.

Data Appendix

Definition of variables

All data are quarterly time series for the sample 1960(1)-1997(4).
Output gap is the deviation of the real GDP from its steady state, approximated by a sequential quadratic trend.
Price inflation is the quarterly growth rate of the total GDP deflator: $\pi_t = 100\,(lnP_t - lnP_{t-1})$.
Wage inflation is the quarterly growth rate of compensation of employees: $w_t = 100\,(lnW_t - lnW_{t-1})$.
Labor income share is the ratio of total compensation and nominal GDP: $mc_t = w_t + h_t - p_t - y_t$, where h_t is (the log of) total employment, and y_t is (the log of) nominal GDP.

Data sources

The data for the United States are from the Bureau of Labor Statistics (BLS) and the Bureau of Economic Analysis (BEA).

Implicit price deflator, non-farm business sector (NFB) = Q.PNF
Employment (persons) (NFB) = M.EEA
Real GDP (NFB) = Q.JQNF
Wage (compensation per hour) (NFB) = Q.JRWSSNF.

Appendix

6.1 Assumptions

We suppose here the same assumptions as in Anatolyev (2005) since we need its results in the following. In particular, parameters are supposed to be strongly identified (Assumption **A2** below). To simplify, let us denote $g_t = g(z_t, \theta)$, $g_t^* = g(z_t, \theta^*)$ and the respective derivative of the function g relative to the parameters as $g_{\theta,t}$ and $g_{\theta,t}^*$.

Assumptions A

- **A1** The sequence z_t is strictly stationary and strongly mixing with mixing coefficients α_j satisfying $\sum_{j=1}^{\infty} j^2 \alpha_j^{1-1/\nu} < \infty$ for some $\nu > 1$.
- **A2** The moment conditions (1) holds for unique $\theta \in int(\Theta)$, where $\Theta \subseteq R^p$ is compact.
- **A3** The function $g(z_t, \theta^*)$ is Borel measurable for all $\theta^* \in \Theta$ and is twice continuously differentiable in θ^* for all $\theta^* \in \Theta$ and for z_t in its support.
- **A4** Form some stationary series d_t with finite $E(d_t^8)$, $\sup_{\theta^* \in \Theta} \max\{\| g_t^* \|, \| g_{\theta,t}^* \|, \| \partial g_{\theta,t}^*/\partial \theta_j \|, \| \partial^2 g_{\theta,t}^*/\partial \theta_j \partial \theta' \| \, \forall j = 1, \cdots, p\} \leq d_t$ and $\max\{\| g_t^* - g_t \|, \| g_{\theta,t}^* - g_{\theta,t} \|, \| \partial g_{\theta,t}^*/\partial \theta_j - \partial g_{\theta,t}/\partial \theta_j \| \, \forall j = 1, \cdots, p\} \leq d_t \| \theta^* - \theta \|$ for all $\theta^* \in \Theta$.
- **A5** The matrices $G = E(g_{\theta,t})$ and $\Omega = \sum_{s=-\infty}^{\infty} E(g_t g_{t-s})$ are of full rank.
- **A6** The kernel function $k(x) : [-b, b] \to [-\bar{k}, \bar{k}]$ for finite b and $\bar{k}$ is symmetric, nonzero at 0, continuous on $(-b, b)$, continuously differentiable on $(-b, b)$ except possibly at a finite number of points, and normalized so that $\int_{-b}^{b} k(x)dx = 1$.
- **A7** $S_T \to \infty$ as $T \to \infty$ and $S_T = o(T^{1/3})$.

For Propositions 1, 2 and 3, we suppose that Assumptions A hold.

Proof of Proposition 1:

The proof is based on Theorem 1 in Robinson (1988) which allows to evaluate the order of magnitude for the stochastic difference between two alternative estimators. The sketch of the proof is closely related to the one in Antoine *et al.* (2007) but with smoothed moment conditions. The proof for the other 3-step (3S-EEL) is similar and omitted for brevity. The p equations corresponding to the FOC for the SEL are

$$f_T(\hat{\theta}_T^{SEL}) = \left[\sum_{t=1}^{T} p_t^{SEL}\left(\hat{\theta}_T^{SEL}\right) G_{tT}\left(\hat{\theta}_T^{SEL}\right)\right]' \left[\widetilde{\Omega}_T^{SEL}(\hat{\theta}_T^{SEL})\right]^{-1} \frac{1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{SEL}) = 0.$$

where $\widetilde{\Omega}_T^{SEL}(\hat{\theta}_T^{SEL}) = S_T \sum_{t=1}^{T} p_t^{SEL}\left(\hat{\theta}_T^{SEL}\right) g_{tT}(\hat{\theta}_T^{SEL}) g_{tT}(\hat{\theta}_T^{SEL})'$ with $p_t^{SEL}\left(\hat{\theta}_T^{SEL}\right) = 1/\left(1 - \hat{\lambda}_T' g_{tT}(\hat{\theta}_T^{SEL})\right)$ is the implied probability at the observation t for the SEL such defined in Smith (2004) with an appropriate definition of $\hat{\lambda}_T$. For the smoothed 3-step estimator with only the weighting matrix evaluated at a second step efficient estimator $\hat{\theta}_T$, the FOC are

$$h_T(\hat{\theta}_T^{3SW}) = \left[\sum_{t=1}^{T} p_t^{SCUE}\left(\hat{\theta}_T^{3SW}\right) G_{tT}\left(\hat{\theta}_T^{3SW}\right)\right]' \left[\widetilde{\Omega}_T^{SCUE}(\hat{\theta}_T)\right]^{-1} \frac{1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{3SW}) = 0.$$

where $\widetilde{\Omega}_T^{SCUE}(\hat{\theta}_T) = S_T \sum_{t=1}^{T} p_t^{SCUE}\left(\hat{\theta}_T\right) g_{tT}(\hat{\theta}_T) g_{tT}(\hat{\theta}_T)'$and $p_t^{SCUE}\left(\hat{\theta}_T\right)$ is defined in eq. (9) .

The objective is to show that $\hat{\theta}_T^{3S} - \hat{\theta}_T^{SEL} = \mathcal{O}_p(T^{-3/2})$. To apply the Theorem 1 in Robinson (1988) two assumptions need to be fulfilled. Assumption A1 in Robinson (1988) is directly verified since $\theta_T^{SEL} = \theta_0 + o_p(1)$. For Assumption A2 in Robinson, we also need that $\theta_T^{3SW} = \theta_0 + o_p(1)$ which is verified. Assumption A2 requires also that the derivative of $h_T(\theta)$ with respect to θ is continuous uniformly in large T with probability arbitrarily close to one in the neighborhood of θ_0 which is guaranteed by Assumption **A4** above.

Under these assumptions, Theorem 1 in Robinson implies that

$$\hat{\theta}_T^{3SW} - \hat{\theta}_T^{SEL} = \mathcal{O}_p\left(\|h_T\left(\hat{\theta}_T^{SEL}\right) - f_T\left(\hat{\theta}_T^{SEL}\right)\|\right).$$

where

$$h_T(\hat{\theta}_T^{SEL}) = \left[\sum_{t=1}^{T} p_t^{SCUE}\left(\hat{\theta}_T^{SEL}\right) G_{tT}\left(\hat{\theta}_T^{SEL}\right)\right]' \left[\widetilde{\Omega}_T^{SCUE}(\hat{\theta}_T)\right]^{-1} \frac{1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{SEL}) = 0.$$

By Therorem 3.1 in Smith (2004), the estimator $\sum_{t=1}^{T} p_t^{SGEL}\left(\hat{\theta}_T^{SGEL}\right) G_{tT}\left(\hat{\theta}_T^{SGEL}\right)$ is an efficient estimator of $G = E\partial g(z_t, \theta_0)/\partial\theta'$ for any SGEL estimator. In particular, the conclusion is valid for the SEL and the smoothed CUE. As mentioned by Smith (2004), this result also holds if the SGEL estimator is replaced by any first order equivalent estimator as the 2-step GMM estimator. This implies that

$$p_t^{SEL}\left(\hat{\theta}_T^{SEL}\right) G_{tT}\left(\hat{\theta}_T^{SEL}\right) = p_t^{SCUE}\left(\hat{\theta}_T\right) G_{tT}\left(\hat{\theta}_T\right) + o_p(1).$$

Consequently,

$$\begin{aligned}\hat{\theta}_T^{3SW} - \hat{\theta}_T^{SEL} &= \mathcal{O}_p\left(\|h_T\left(\hat{\theta}_T^{SEL}\right) - f_T\left(\hat{\theta}_T^{SEL}\right)\|\right) \\ &\leq \mathcal{O}_p\left(\left\|\left[\sum_{t=1}^{T} p_t^{SEL}\left(\hat{\theta}_T^{SEL}\right) G_{tT}\left(\hat{\theta}_T^{SEL}\right)\right]'\right\| \|\widetilde{\Omega}_T^{SCUE}(\hat{\theta}_T)^{-1} - \widetilde{\Omega}_T^{SEL}(\hat{\theta}_T^{SEL})^{-1}\| \left\|\frac{1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{SEL})\right\|\right)\end{aligned}$$

Since $\frac{1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T^{SEL}) = \mathcal{O}_p(1/\sqrt{T})$ and $\sum_{t=1}^{T} p_t^{SEL}\left(\hat{\theta}_T^{SEL}\right) G_{tT}\left(\hat{\theta}_T^{SEL}\right) \xrightarrow{p} G$, we only need to show that:

$$\left\|\widetilde{\Omega}^{SCUE}(\hat{\theta}_T)^{-1} - \widetilde{\Omega}^{SEL}(\hat{\theta}_T^{SEL})^{-1}\right\| = \mathcal{O}_p(1/T). \tag{14}$$

Consider the smoothed CUE, namely $\hat{\theta}_T^{SCUE}$, by virtue of the triangular inequality

$$\left\|\widetilde{\Omega}^{SCUE}(\hat{\theta}_T) - \widetilde{\Omega}^{SEL}(\hat{\theta}_T^{SEL})\right\| \leq \left\|\widetilde{\Omega}^{SCUE}(\hat{\theta}_T) - \widetilde{\Omega}^{SCUE}(\hat{\theta}_T^{SCUE})\right\| + \left\|\widetilde{\Omega}^{SCUE}(\hat{\theta}_T^{SCUE}) - \widetilde{\Omega}^{SEL}(\hat{\theta}_T^{SEL})\right\|.$$

The first expression at the right hand side is $\mathcal{O}_p(1/T)$ by an usual Taylor expansion and $\hat{\theta}_T - \hat{\theta}_T^{CUE} = \mathcal{O}_p(1/T)$. The second expression is also $\mathcal{O}_p(1/T)$ by a direct implication of Theorem 3.1 in Smith (2004) for efficient moment estimators of variance-covariance matrix of the moment conditions by CUE and SEL. The result follows by noticing that $M^{-1} - N^{-1} = M^{-1}(N - M)N^{-1}$.

Proof of Proposition 2:
Theorem 1 in Anatolyev (2005) provides the asymptotic bias of the SEL estimator. By Proposition 1, the smoothed 3-step estimator achieves the same higher order efficiency, e.g. the asymptotic bias of this estimator is the same as the one for the SEL estimator up to an order $\mathcal{O}_p(T^{-3/2})$. The first term appearing in the asymptotic bias of the SEL estimator (Theorem 1 in Anatolyev (2005)) is removed for instance by the use of the uniform kernel proposed by Kitamura and Stutzer (1997). The asymptotic bias at order T^{-1} of the 3SW-EEL estimator is then given by: $B_{G\Xi g} + B_{\partial^2 g}$.

Proof of Proposition 3:
The implied probabilities evaluated at the (bias-corrected) smoothed 3S-EEL (or smoothed 3SW-EEL) estimator are defined by Definition 1 as:

$$p_t^{SCUE}(\hat{\theta}_T^{3S}) = \frac{1}{T} - \frac{1}{T}\frac{S_T}{k_2}\left[g_{tT}(\hat{\theta}_T^{3S}) - \overline{g}_T(\hat{\theta}_T^{3S})\right]'\hat{\Omega}_T(\hat{\theta}_T^{3S})^{-1}\overline{g}_T(\hat{\theta}_T^{3S})$$

or equivalently:

$$p_t^{SCUE}(\hat{\theta}_T^{3S}) = \frac{1}{T} - \frac{1}{T}\frac{S_T}{k_2}\overline{g}_T(\hat{\theta}_T^{3S})'\hat{\Omega}_T(\hat{\theta}_T^{3S})^{-1}\left[g_{tT}(\hat{\theta}_T^{3S}) - \overline{g}_T(\hat{\theta}_T^{3S})\right].$$

Now we compute $\sum_{t=1}^{T}\left(p_t^{SCUE}(\hat{\theta}_T^{3S}) - \frac{1}{T}\right)^2$, this yields:

$$\begin{aligned}\sum_{t=1}^{T}\left(p_t^{SCUE}(\hat{\theta}_T^{3S}) - \frac{1}{T}\right)^2 &= \frac{1}{T}\frac{S_T}{k_2}\overline{g}_T(\hat{\theta}_T^{3S})'\hat{\Omega}_T(\hat{\theta}_T^{3S})^{-1}\frac{1}{T}\frac{S_T}{k_2}\sum_{t=1}^{T}\left[g_{tT}(\hat{\theta}_T^{3S}) - \overline{g}_T(\hat{\theta}_T^{3S})\right] \\ &\quad\times\left[g_{tT}(\hat{\theta}_T^{3S}) - \overline{g}_T(\hat{\theta}_T^{3S})\right]'\hat{\Omega}_T(\hat{\theta}_T^{3S})^{-1}\overline{g}_T(\hat{\theta}_T^{3S})\end{aligned}$$

An consistent estimator of Ω in mean deviation is obtained by:

$$\frac{1}{T}\frac{S_T}{k_2}\sum_{t=1}^{T}\left[g_{tT}(\hat{\theta}_T^{3S}) - \overline{g}_T(\hat{\theta}_T^{3S})\right]\left[g_{tT}(\hat{\theta}_T^{3S}) - \overline{g}_T(\hat{\theta}_T^{3S})\right]'.$$

This gives that:

$$\frac{1}{T}\sum_{t=1}^{T}\left(Tp_t^{SCUE}(\hat{\theta}_T^{3S}) - 1\right)^2 = \frac{S_T}{k_2}\overline{g}_T(\hat{\theta}_T^{3S})'\hat{\Omega}_T(\hat{\theta}_T^{3S})^{-1}\overline{g}_T(\hat{\theta}_T^{3S})$$

where the right hand side expression corresponds to the GMM criteria evaluated at the bias-corrected smoothed 3S-EEL estimator. Multiplying by T, this yields:

$$\frac{k_2}{k_1^2 S_T}\sum_{t=1}^{T}\left(Tp_t^{SCUE}(\hat{\theta}_T^{3S}) - 1\right)^2 = T\overline{g}_T(\hat{\theta}_T^{3S})'\hat{\Omega}_T(\hat{\theta}_T^{3S})^{-1}\overline{g}_T(\hat{\theta}_T^{3S})/k_1^2.$$

Hence, the IPS_T statistic is numerically equivalent to the J-statistic but for smoothed moment conditions with a centered weighting matrix (see Smith 2005). This statistic is asymptotically equivalent to the standard optimal GMM statistic J by Theorem 4.1 in Smith (2004) and Theorem 3.2 in Smith (2005). The IPS_T statistic is then asymptotically distributed as $\chi^2(q-p)$.

References

Anatolyev, S., 2005, "GMM, GEL, Serial Correlation, and Asymptotic Bias", *Econometrica* 73: 983-1002.

Anderson, T.W., and N. Kunitomo, 2005, "A New Light from Old Wisdoms: Alternative Estimation Methods of Simultaneous Equations and Microeconometric Models ", CIRJE Discussion Papers.

Andrews, D.W.K., 1991, "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica* 59(3): 817-858.

Antoine, B, Bonnal, H, and E. Renault, 2007, "On the Efficient Use of the Informational Content of Estimating Equations: Implied Probabilities and Euclidean Empirical Likelihood", *Journal of Econometrics* 127(2): 461-487.

Back, K. and D.P. Brown, 1993, "Implied Probabilities in GMM Estimators" *Econometrica* 61: 971-75.

Baggerly, K.A., 1998, "Empirical Likelihood as a Goodness-of-Fit Measure" *Biometrika* 85: 535-47.

Bonnal, H. and E. Renault, 2001, "Minimal Chi-Square Estimation with Conditional Moment Restrictions, " Document de Travail, CESG, September 2001.

Brown, B.W., and W.K. Newey, 1998, "Efficient Semiparametric Estimation of Expectations ", *Econometrica* 66(2): 453-464.

Buiter, W. and I. Jewitt, 1989, "Staggered Wage Setting with Real Wage Relativities: Variations on a Theme of Taylor ", in W. Buiter (Ed.), *Macroeconomic Theory and Stabilization Policy*, 183-199, University of Michigan Press.

Calvo, G., 1983, "Staggered Prices in a Utility-Maximizing Framework", *Journal of Monetary Economics* 12: 383-98.

Chamberlain, G., 1987, "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions", *Journal of Econometrics* 34: 305-334.

Cumby, R.E. and J. Huizinga, 1992, "Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressors", *Econometrica* 60(1): 185-195.

Dufour, J.M., 1997,"Some Impossibility Theorems in Econometrics, with Applications to Structural and Dynamic Models", *Econometrica* 65: 1365-1389.

—, 2003, "Identification, Weak Instruments and Statistical Inference in Econometrics", *Canadian Journal of Economics* 36(4), 767-808.

Dufour, J.M., L. Khalaf, and M. Kichian, 2005, "Estimating New Keynesian Phillips Curves using Optimal Instruments" *Journal of Economic Dynamics and Control*, Volume 30(9-10): 1707-1727

Galí, J. and M. Gertler, 1999, "Inflation Dynamics: A Structural Econometrics Analysis", *Journal of Monetary Economics* 44: 195-222.

Galí, J., M. Gertler, and J. D. Lopez-Salido, 2001, "European Inflation Dynamics" *European Economic Review* 45: 1237-70.

—, 2005, "Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve " *Journal of Monetary Economics* 52(6): 1107-1118

Guay, A. and F. Pelgrin, 2004, "The U.S. New Keynesian Phillips Curve: An Empirical Assessment", Working paper, Bank of Canada.

—, 2007, "Efficient Inference Using Implied Probabilities under Weak Identification ", mimeo.

Guggenberger, P., 2006, "Finite Sample Evidence Suggesting a Heavy Tail Problem of the Generalized Empirical Likelihood Estimator", *Econometric Reviews*, forthcoming.

—, 2005, "Monte-Carlo Evidence Suggesting a no Moment Problem of the Continuous Updating Estimator", *Economics Bulletin* 3: 1-6.

Guggenberger, P. and J. Hahn, 2005, "Finite Sample Properties of the Two-Step Empirical Likelihood Estimator ", *Econometric Reviews* 24(3): 247-263.

Guggenberger, P. and R.S. Smith, 2007, "Generalized Empirical Likelihood Tests in Time Series Models with Potential Identification Failure", *Journal of Econometrics*, forthcoming.

Hall, A.R., 2000, "Covariance Matrix Estimation and the Power of the Overidentifying Restrictions", *Econometrica* 68: 1517-28.

Hansen, L.P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica* 50: 1029-54.

Hansen, L.P., J. Heaton, and A. Yaron, 1996, "Finite-Sample Properties of Some Alternative GMM Estimators", *Journal of Business and Economic Statistics* 14: 262-80.

Imbens, G., Spady, R. and P. Johnson, 1998, "Information Theoretic Approaches to Inference in Moment Condition Models", *Econometrica* 66(2): 333-357.

Jondeau, E. and H. Le Bihan, 2007, "ML vs GMM Estimates of Hybrid Macroeconomic Models (with an Application to the New Phillips Curve)", *Journal of Econometrics*, forthcoming.

Kitamura, Y. and M. Stutzer, 1997, "An Information-Theoretic Alternative to Generalized Method of Moments Estimation", *Econometrica* 65: 861-74.

Kleibergen, F., 2002, "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression", *Econometrica* 70(5): 1781-1805.

—, 2005, "Testing Parameters in GMM without Assuming that they are Identified", *Econometrica* 73(4): 1103-1123.

—, 2007, "Generalizing Weak Instrument Robust IV Statistics Towards Multiple Parameters, Unrestricted Covariance Matrices and Identification Statistics", *Journal of Econometrics* 139(1): 181-216.

Kurmann, A., 2005, "Quantifying the Uncertainty about the Fit of a New Keynesian Pricing Model", *Journal of Monetary Economics* 52(6): 1119-1134.

Lindé, J., 2005, "Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach", *Journal of Monetary Economics* 52(6): 1135–1142.

Ma, A. 2002. "GMM Estimation of the New Phillips Curve", *Economic Letters* 76(3): 411-17.

Mavroeidis, S., 2004, "Identification and Misspecification Issues in Forward Looking Monetary Models", *Oxford Bulletin of Economics and Statistics* 66(s1): 609-36.

—, 2005, "Identification Issues in Forward-Looking Models Estimated by GMM, with an Application to the Phillips Curve", *Journal of Money, Credit, and Banking*, vol. 37(3): 421-448.

—, 2007, "Testing the New Keynesian Phillips Curve Without Assuming Identification", mimeo.

Nason, J.M. and G.W. Smith, 2005, "Identifying the New Keynesian Phillips Curve", Working Paper 2005-01, Federal Reserve Bank of Atlanta.

Newey, W.K. and R.J. Smith, 2004, "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators", *Econometrica* 72: 219-55.

Newey, W.K. and K. West, 1994, "Automatic Lag Selection in Covariance Matrix Estimation", *Review of Economic Studies* 61: 631-53.

—, 1987, "A Simple, Positive Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica* 55(3): 703–708.

Owen, A.B., 1990, "Empirical Likelihood Ratio Confidence Regions", *The Annals of Statistics* 18: 90-120.

—, 1991, "Empirical Likelihood for Linear Models", *The Annals of Statistics* 19: 1725:1747.

—, 2001, *Empirical Likelihood*, Monographs on Statistics and Applied Probability 92, Chapman & Hall.

Qin, J. and J. Lawless, 1994, "Empirical Likelihood and General Estimating Equations", *Annals of Statistics* 22(1): 300-325.

Ramalho, J.J.S. and R.J. Smith, 2005, "Goodness of Fit Test for Moment Condition Models", Working Paper, Universidade de Évora, No 2005-05.

Robinson, P.M., 1988, "The Stochastic Difference between Econometric Statistics", *Econometrica* 56: 531-548.

Rudd, J. and K. Whelan, 2005, "New Tests of the New Keynesian Phillips Curve", *Journal of Monetary Economics* 52(6): 1167–1181.

—, 2006, "Can Rational Expectations Sticky-Price Models Explain Inflation Dynamics?", *American Economic Review* 96(1): 303-320.

Schennach, S.M., 2007, "Point Estimation with Exponentially Tilted Empirical Likelihood", *The Annals of Statistics* 35: 634-672.

Smith, R.J., 2000, "Empirical Likelihood Estimation and Inference", in M. Salmon and P. Marriott (Eds), *Applications of Differential Geometry to Econometrics*, chapter 4, Cambridge University Press, 119-150.

—, 2004, "GEL Criteria for Moment Condition Models, " mimeo.

—, 2005, "Automatic Positive Semi-Definite HAC Covariance Matrix and GMM Estimation", *Econometric Theory* 21(1): 158-170.

—, 2007, "Efficient Information Theoretic Inference for Conditional Moment Restrictions", *Journal of Econometrics*, 138(2): 430-460.

Stock, J. and W. Wright, 2000, "GMM with Weak Identification", *Econometrica* 68: 1055-96.

West, K., 1997, "Another Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator" *Journal of Econometrics* 76: 171-91.

Table 1: Finite sample properties in DGP1 ($\rho_1 = .9(1-\rho_2)$, $\rho_2 = -.65$ and $T = 160$)

				$\gamma_f = .591$		$\gamma_b = .378$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	12.10	-.0887	.1204	.0003	.0429	.0085	.0120
3S-EEL	8	0	12.10	-.0796	.1202	-.0035	.0461	.0081	.0119
3SW-EEL	8	0	12.10	-.0782	.1290	-.0038	.0486	.0080	.0122
3SWc-EEL	8	0	12.10	-.0737	.1198	-.0002	.0468	.0068	.0111
CUE	8	0	12.10	-.0769	.1323	-.0024	.0548	-.0251	1.5958
2S-GMM	16	0	12.10	-.1121	.1350	.0062	.0431	.0101	.0133
3S-EEL	16	0	12.10	-.0989	.1290	.0044	.0472	.0089	.0128
3SW-EEL	16	0	12.10	-.0978	.1486	.0041	.0530	.0087	.0139
3SWc-EEL	16	0	12.10	-.0945	.1438	.0067	.0523	.0079	.0132
CUE	16	0	12.10	-.1020	.1793	-.0012	.0988	-1.1696	73.658
2S-GMM	24	0	12.10	-.1295	.1480	.0105	.0431	.0114	.0145
3S-EEL	24	0	12.10	-.1165	.1414	.0120	.0479	.0098	.0137
3SW-EEL	24	0	12.10	-.1160	.1631	.0132	.0557	.0095	.0151
3SWc-EEL	24	0	12.10	-.1137	.1604	.0150	.0558	.0089	.0147
CUE	24	0	12.10	-.1534	.2725	.0175	.1870	-5.5305	148.74
2S-GMM	8	.5	8.87	-.0899	.1227	-.0086	.0519	.0111	.0150
3S-EEL	8	.5	8.87	-.0765	.1194	-.0094	.0559	.0096	.0138
3SW-EEL	8	.5	8.87	-.0744	.1290	-.0085	.0587	.0092	.0139
3SWc-EEL	8	.5	8.87	-.0716	.1205	-.0043	.0556	.0081	.0129
CUE	8	.5	8.87	-.0752	.1399	-.0073	.0655	.0068	.1243
2S-GMM	16	.5	8.87	-.1219	.1437	-.0101	.0525	.0152	.0190
3S-EEL	16	.5	8.87	-.1048	.1342	-.0051	.0553	.0123	.0168
3SW-EEL	16	.5	8.87	-.0993	.1499	-.0008	.0636	.0105	.0163
3SWc-EEL	16	.5	8.87	-.0965	.1457	.0035	.0625	.0097	.0157
CUE	16	.5	8.87	-.0982	.1919	-.0052	.1163	-.4380	21.513
2S-GMM	24	.5	8.87	-.1430	.1602	-.0164	.0542	.0190	.0225
3S-EEL	24	.5	8.87	-.1264	.1500	-.0081	.0560	.0157	.0201
3SW-EEL	24	.5	8.87	-.1210	.1668	.0026	.0655	.0132	.0196
3SWc-EEL	24	.5	8.87	-.1190	.1643	.0045	.0650	.0126	.0191
CUE	24	.5	8.87	-.1374	.2890	-.0030	.2112	-282.425	1695.4
2S-GMM	8	-.5	16.35	-.0905	.1239	-.0023	.0409	.0079	.0114
3S-EEL	8	-.5	16.35	-.0850	.1257	-.0074	.0438	.0081	.0119
3SW-EEL	8	-.5	16.35	-.0831	.1311	-.0087	.0451	.0081	.0123
3SWc-EEL	8	-.5	16.35	-.0768	.1214	-.0051	.0435	.0069	.0111
CUE	8	-.5	16.35	-.0853	.1360	-.0056	.0481	.0096	.0783
2S-GMM	16	-.5	16.35	-.1101	.1370	.0051	.0414	.0084	.0118
3S-EEL	16	-.5	16.35	-.1019	.1356	.0011	.0458	.0083	.0121
3SW-EEL	16	-.5	16.35	-.1002	.1513	-.0027	.0491	.0087	.0136
3SWc-EEL	16	-.5	16.35	-.0958	.1457	-.0010	.0482	.0078	.0129
CUE	16	-.5	16.35	-.1121	.1862	-.0069	.0937	-.0247	1.5321
2S-GMM	24	-.5	16.35	-.1216	.1443	.0119	.0417	.0085	.0118
3S-EEL	24	-.5	16.35	-.1128	.1422	.0109	.0463	.0079	.0119
3SW-EEL	24	-.5	16.35	-.1128	.1622	.0076	.0516	.0084	.0139
3SWc-EEL	24	-.5	16.35	-.1100	.1590	.0093	.0515	.0078	.0136
CUE	24	-.5	16.35	-.1509	.2658	.0079	.1904	-1.1256	75.098

Table 2: Large sample properties in DGP1 ($\rho_1 = .9(1-\rho_2)$, $\rho_2 = -.65$ and $T = 500$)

				$\gamma_f = .591$		$\gamma_b = .378$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	37.48	-.0301	.0492	.0009	.0215	.0027	.0047
3S-EEL	8	0	37.48	-.0242	.0475	-.0016	.0221	.0025	.0046
3SW-EEL	8	0	37.48	-.0239	.0475	-.0016	.0221	.0025	.0046
3SWc-EEL	8	0	37.48	-.0238	.0460	-.0000	.0218	.0021	.0043
CUE	8	0	37.48	-.0224	.0467	-.0004	.0220	.0022	.0045
2S-GMM	16	0	37.48	-.0421	.0568	.0050	.0219	.0033	.0051
3S-EEL	16	0	37.48	-.0303	.0506	.0015	.0223	.0026	.0047
3SW-EEL	16	0	37.48	-.0291	.0504	.0014	.0224	.0025	.0047
3SWc-EEL	16	0	37.48	-.0289	.0494	.0027	.0222	.0022	.0045
CUE	16	0	37.48	-.0232	.0500	.0013	.0232	.0020	.0046
2S-GMM	24	0	37.48	-.0530	.0653	.0080	.0228	.0039	.0057
3S-EEL	24	0	37.48	-.0379	.0557	.0051	.0232	.0028	.0050
3SW-EEL	24	0	37.48	-.0361	.0551	.0050	.0233	.0027	.0050
3SWc-EEL	24	0	37.48	-.0357	.0539	.0061	.0234	.0024	.0048
CUE	24	0	37.48	-.0247	.0544	.0028	.0249	.0020	.0052
2S-GMM	8	.5	27.50	-.0319	.0507	-.0022	.0248	.0037	.0059
3S-EEL	8	.5	27.50	-.0243	.0477	-.0038	.0256	.0030	.0055
3SW-EEL	8	.5	27.50	-.0238	.0475	-.0038	.0256	.0030	.0055
3SWc-EEL	8	.5	27.50	-.0248	.0462	-.0018	.0248	.0027	.0052
CUE	8	.5	27.50	-.0215	.0465	-.0028	.0256	.0027	.0054
2S-GMM	16	.5	27.50	-.0474	.0607	.0013	.0243	.0048	.0068
3S-EEL	16	.5	27.50	-.0324	.0517	-.0003	.0251	.0034	.0058
3SW-EEL	16	.5	27.50	-.0309	.0509	-.0004	.0252	.0032	.0057
3SWc-EEL	16	.5	27.50	-.0314	.0498	.0012	.0248	.0029	.0055
CUE	16	.5	27.50	-.0217	.0486	-.0006	.0269	.0023	.0055
2S-GMM	24	.5	27.50	-.0616	.0727	.0030	.0245	.0062	.0081
3S-EEL	24	.5	27.50	-.0431	.0591	.0033	.0256	.0040	.0065
3SW-EEL	24	.5	27.50	-.0411	.0580	.0033	.0258	.0038	.0064
3SWc-EEL	24	.5	27.50	-.0411	.0571	.0046	.0256	.0035	.0061
CUE	24	.5	27.50	-.0230	.0550	.0020	.0290	.0021	.0059
2S-GMM	8	-.5	50.66	-.0311	.0513	-.0001	.0208	.0026	.0045
3S-EEL	8	-.5	50.66	-.0268	.0502	-.0026	.0214	.0026	.0045
3SW-EEL	8	-.5	50.66	-.0266	.0503	-.0026	.0214	.0025	.0045
3SWc-EEL	8	-.5	50.66	-.0254	.0484	-.0013	.0211	.0022	.0042
CUE	8	-.5	50.66	-.0259	.0499	-.0013	.0212	.0023	.0044
2S-GMM	16	-.5	50.66	-.0398	.0573	.0034	.0210	.0028	.0047
3S-EEL	16	-.5	50.66	-.0312	.0537	-.0002	.0215	.0026	.0047
3SW-EEL	16	-.5	50.66	-.0303	.0536	-.0003	.0216	.0025	.0046
3SWc-EEL	16	-.5	50.66	-.0293	.0520	.0009	.0214	.0022	.0044
CUE	16	-.5	50.66	-.0275	.0539	.0000	.0221	.0023	.0047
2S-GMM	24	-.5	50.66	-.0476	.0625	.0062	.0219	.0031	.0049
3S-EEL	24	-.5	50.66	-.0361	.0569	.0028	.0225	.0026	.0048
3SW-EEL	24	-.5	50.66	-.0348	.0566	.0026	.0226	.0025	.0048
3SWc-EEL	24	-.5	50.66	-.0338	.0551	.0037	.0226	.0022	.0045
CUE	24	-.5	50.66	-.0302	.0582	.0013	.0238	.0026	.0053

Table 3: Finite sample properties in DGP2 ($\rho_1 = .9(1-\rho_2)$, $\rho_2 = -.65$ and $T = 160$)

				$\gamma_f = .85$		$\gamma_b = .10$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	15.74	-.1910	.2225	.0295	.0709	.0029	.0055
3S-EEL	8	0	15.74	-.1431	.1850	.0213	.0699	.0021	.0051
3SW-EEL	8	0	15.74	-.1274	.1854	.0194	.0710	.0018	.0051
3SWc-EEL	8	0	15.74	-.1377	.1879	.0269	.0727	.0015	.0049
CUE	8	0	15.74	-.1295	.1792	.0241	.0741	-.0010	.1758
2S-GMM	16	0	15.74	-.2860	.3101	.0416	.0786	.0048	.0071
3S-EEL	16	0	15.74	-.2336	.2682	.0445	.0840	.0031	.0062
3SW-EEL	16	0	15.74	-.2008	.2687	.0449	.0903	.0022	.0062
3SWc-EEL	16	0	15.74	-.2032	.2684	.0497	.0925	.0018	.0061
CUE	16	0	15.74	-.1882	.2787	.0488	.1438	-1.5817	45.889
2S-GMM	24	0	15.74	-.3502	.3684	.0442	.0810	.0066	.0087
3S-EEL	24	0	15.74	-.3071	.3327	.0533	.0903	.0048	.0076
3SW-EEL	24	0	15.74	-.2735	.3272	.0608	.1019	.0033	.0075
3SWc-EEL	24	0	15.74	-.2730	.3259	.0641	.1038	.0030	.0073
CUE	24	0	15.74	-.2634	.3857	.0788	.2297	-8.7313	135.99
2S-GMM	8	.5	12.39	-.2309	.2674	.0582	.0967	.0017	.0055
3S-EEL	8	.5	12.39	-.1817	.2315	.0503	.0962	.0008	.0052
3SW-EEL	8	.5	12.39	-.1812	.2519	.0529	.1059	.0005	.0056
3SWc-EEL	8	.5	12.39	-.1917	.2531	.0625	.1097	-.0001	.0055
CUE	8	.5	12.39	-.1679	.2401	.0519	.1057	.1301	5.4078
2S-GMM	16	.5	12.39	-.3151	.3394	.0647	.1009	.0041	.0073
3S-EEL	16	.5	12.39	-.2618	.2977	.0795	.1092	.0019	.0064
3SW-EEL	16	.5	12.39	-.2532	.3216	.0801	.1269	.0006	.0066
3SWc-EEL	16	.5	12.39	-.2548	.3207	.0851	.1297	.0003	.0066
CUE	16	.5	12.39	-.2139	.2978	.0695	.0998	.0350	.0490
2S-GMM	24	.5	12.39	-.3756	.3938	.0615	.0990	.0066	.0092
3S-EEL	24	.5	12.39	-.3357	.3616	.0746	.1117	.0043	.0080
3SW-EEL	24	.5	12.39	-.3173	.3682	.0924	.1357	.0022	.0078
3SWc-EEL	24	.5	12.39	-.3162	.3663	.0953	.1376	.0020	.0077
CUE	24	.5	12.39	-.3028	.4305	.1295	.2788	-174.97	9135.9
2S-GMM	8	-.5	22.68	-.1610	.1878	.0125	.0612	.0032	.0056
3S-EEL	8	-.5	22.68	-.1190	.1540	.0077	.0617	.0026	.0051
3SW-EEL	8	-.5	22.68	-.0960	.1410	.0074	.0627	.0021	.0049
3SWc-EEL	8	-.5	22.68	-.1076	.1439	.0129	.0633	.0018	.0047
CUE	8	-.5	22.68	-.1085	.1461	.0139	.0654	.0022	.0051
2S-GMM	16	-.5	22.68	-.2540	.2756	.0217	.0658	.0048	.0069
3S-EEL	16	-.5	22.68	-.2012	.2300	.0242	.0698	.0035	.0061
3SW-EEL	16	-.5	22.68	-.1483	.2004	.0244	.0737	.0023	.0057
3SWc-EEL	16	-.5	22.68	-.1508	.2005	.0286	.0750	.0020	.0056
CUE	16	-.5	22.68	-.1578	.2326	.0279	.1173	-.0815	1.9622
2S-GMM	24	-.5	22.68	-.3183	.3367	.0245	.0673	.0061	.0080
3S-EEL	24	-.5	22.68	-.2731	.2993	.0334	.0752	.0046	.0071
3SW-EEL	24	-.5	22.68	-.2195	.2723	.0382	.0839	.0031	.0069
3SWc-EEL	24	-.5	22.68	-.2193	.2717	.0414	.0853	.0029	.0068
CUE	24	-.5	22.68	-.2315	.3427	.0497	.1961	-5.1795	134.35

Table 4: **Large sample properties in DGP2** ($\rho_1 = .9(1 - \rho_2)$, $\rho_2 = -.65$ **and** $T = 500$)

				$\gamma_f = .85$		$\gamma_b = .10$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	48.77	-.0706	.0914	.0062	.0461	.0012	.0030
3S-EEL	8	0	48.77	-.0442	.0755	-.0007	.0470	.0010	.0029
3SW-EEL	8	0	48.77	-.0430	.0752	-.0007	.0471	.0010	.0029
3SWc-EEL	8	0	48.77	-.0498	.0775	.0026	.0468	.0009	.0028
CUE	8	0	48.77	-.0442	.0759	.0022	.0475	.0009	.0029
2S-GMM	16	0	48.77	-.1144	.1296	.0167	.0489	.0016	.0032
3S-EEL	16	0	48.77	-.0659	.0921	.0080	.0496	.0009	.0030
3SW-EEL	16	0	48.77	-.0618	.0899	.0079	.0498	.0008	.0030
3SWc-EEL	16	0	48.77	-.0670	.0921	.0107	.0501	.0007	.0029
CUE	16	0	48.77	-.0430	.0807	.0061	.0519	.0006	.0031
2S-GMM	24	0	48.77	-.1552	.1683	.0241	.0526	.0022	.0037
3S-EEL	24	0	48.77	-.0980	.1194	.0189	.0542	.0010	.0032
3SW-EEL	24	0	48.77	-.0920	.1152	.0187	.0546	.0009	.0032
3SWc-EEL	24	0	48.77	-.0956	.1170	.0211	.0552	.0007	.0031
CUE	24	0	48.77	-.0462	.0896	.0111	.0586	-.0017	.0931
2S-GMM	8	.5	38.39	-.0864	.1089	.0197	.0555	.0006	.0031
3S-EEL	8	.5	38.39	-.0531	.0864	.0090	.0541	.0005	.0032
3SW-EEL	8	.5	38.39	-.0516	.0874	.0088	.0544	.0005	.0032
3SWc-EEL	8	.5	38.39	-.0622	.0912	.0142	.0548	.0003	.0031
CUE	8	.5	38.39	-.0505	.0847	.0110	.0553	.0004	.0032
2S-GMM	16	.5	38.39	-.1444	.1612	.0363	.0637	.0010	.0033
3S-EEL	16	.5	38.39	-.0871	.1134	.0238	.0607	.0002	.0033
3SW-EEL	16	.5	38.39	-.0810	.1088	.0228	.0607	.0002	.0033
3SWc-EEL	16	.5	38.39	-.0883	.1122	.0271	.0618	.0000	.0033
CUE	16	.5	38.39	-.0519	.0955	.0154	.0625	.0009	.0035
2S-GMM	24	.5	38.39	-.1930	.2073	.0466	.0709	.0016	.0038
3S-EEL	24	.5	38.39	-.1299	.1510	.0396	.0701	.0002	.0035
3SW-EEL	24	.5	38.39	-.1229	.1451	.0387	.0700	.0002	.0035
3SWc-EEL	24	.5	38.39	-.1270	.1470	.0418	.0714	.0000	.0035
CUE	24	.5	38.39	-.0605	.1096	.0232	.0716	-.0077	.2675
2S-GMM	8	-.5	70.27	-.0598	.0821	-.0025	.0433	.0014	.0029
3S-EEL	8	-.5	70.27	-.0384	.0701	-.0068	.0445	.0012	.0028
3SW-EEL	8	-.5	70.27	-.0375	.0699	-.0068	.0445	.0012	.0028
3SWc-EEL	8	-.5	70.27	-.0418	.0711	-.0046	.0441	.0010	.0027
CUE	8	-.5	70.27	-.0397	.0709	-.0041	.0445	.0011	.0028
2S-GMM	16	-.5	70.27	-.0948	.1108	.0038	.0441	.0018	.0032
3S-EEL	16	-.5	70.27	-.0544	.0817	-.0012	.0462	.0012	.0030
3SW-EEL	16	-.5	70.27	-.0514	.0803	-.0011	.0464	.0011	.0029
3SWc-EEL	16	-.5	70.27	-.0548	.0815	.0008	.0462	.0010	.0029
CUE	16	-.5	70.27	-.0418	.0763	-.0005	.0477	.0009	.0029
2S-GMM	24	-.5	70.27	-.1267	.1399	.0082	.0450	.0023	.0036
3S-EEL	24	-.5	70.27	-.0767	.0997	.0055	.0477	.0013	.0031
3SW-EEL	24	-.5	70.27	-.0720	.0969	.0058	.0480	.0012	.0030
3SWc-EEL	24	-.5	70.27	-.0748	.0981	.0075	.0481	.0011	.0030
CUE	24	-.5	70.27	-.0457	.0851	.0026	.0505	.0012	.0129

Table 5: Finite sample properties in DGP3 ($\rho_1 = .9(1-\rho_2)$, $\rho_2 = -.65$ and $T = 160$)

				$\gamma_f = .10$		$\gamma_b = .85$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	.66	.2859	.2962	-.2519	.2611	-.0085	.0091
3S-EEL	8	0	.66	.2512	.2722	-.2228	.2411	-.0073	.0083
3SW-EEL	8	0	.66	.2219	.2558	-.1978	.2270	-.0063	.0079
3SWc-EEL	8	0	.66	.2596	.2791	-.2298	.2461	-.0078	.0087
CUE	8	0	.66	.2212	.2623	-.1981	.2342	-.0010	.1722
2S-GMM	16	0	.66	.3276	.3322	-.2887	.2927	-.0100	.0103
3S-EEL	16	0	.66	.3092	.3178	-.2733	.2809	-.0092	.0097
3SW-EEL	16	0	.66	.2702	.2927	-.2402	.2595	-.0079	.0090
3SWc-EEL	16	0	.66	.2783	.2992	-.2474	.2646	-.0083	.0093
CUE	16	0	.66	.2240	.2835	-.2370	.2998	-.1997	4.3342
2S-GMM	24	0	.66	.3504	.3530	-.3091	.3115	-.0108	.0110
3S-EEL	24	0	.66	.3412	.3455	-.3017	.3057	-.0104	.0104
3SW-EEL	24	0	.66	.3119	.3252	-.2774	.2886	-.0094	.0094
3SWc-EEL	24	0	.66	.3144	.3271	-.2791	.2899	-.0095	.0101
CUE	24	0	.66	.2173	.3065	-.3096	.4027	-.3092	37.803
2S-GMM	8	.5	.57	.2808	.2900	-.2472	.2712	-.0083	.0088
3S-EEL	8	.5	.57	.2530	.2712	-.2224	.2392	-.0074	.0083
3SW-EEL	8	.5	.57	.2296	.2579	-.2035	.2277	-.0067	.0079
3SWc-EEL	8	.5	.57	.2628	.2795	-.2312	.2457	-.0078	.0086
CUE	8	.5	.57	.2288	.2647	-.2039	.2356	-.0064	.0109
2S-GMM	16	.5	.57	.3208	.3249	-.2823	.2859	-.0094	.0097
3S-EEL	16	.5	.57	.3091	.3162	-.2723	.2784	-.0091	.0095
3SW-EEL	16	.5	.57	.2770	.2962	-.2445	.2613	-.0081	.0090
3SWc-EEL	16	.5	.57	.2841	.3011	-.2505	.2651	-.0084	.0092
CUE	16	.5	.57	.2821	.3362	-.2505	.3094	-.0048	.0343
2S-GMM	24	.5	.57	.3358	.3383	-.2960	.2983	-.0097	.0100
3S-EEL	24	.5	.57	.3301	.3340	-.2910	.2945	-.0096	.0099
3SW-EEL	24	.5	.57	.3090	.3198	-.2731	.2825	-.0089	.0095
3SWc-EEL	24	.5	.57	.3110	.3214	-.2746	.2836	-.0090	.0096
CUE	24	.5	.57	.2296	.3245	-.3240	.4210	-2.2364	106.57
2S-GMM	8	-.5	.76	.2926	.3044	-.2594	.2701	-.0087	.0093
3S-EEL	8	-.5	.76	.2504	.2748	-.2244	.2458	-.0072	.0084
3SW-EEL	8	-.5	.76	.2189	.2571	-.1980	.2307	-.0062	.0079
3SWc-EEL	8	-.5	.76	.2625	.2846	-.2335	.2528	-.0080	.0088
CUE	8	-.5	.76	.2174	.2618	-.1970	.2353	-.0061	.0086
2S-GMM	16	-.5	.76	.3456	.3510	-.3070	.3118	-.0108	.0111
3S-EEL	16	-.5	.76	.3216	.3315	-.2872	.2961	-.0099	.0104
3SW-EEL	16	-.5	.76	.2716	.3001	-.2452	.2694	-.0082	.0095
3SWc-EEL	16	-.5	.76	.2826	.3076	-.2536	.2750	-.0087	.0099
CUE	16	-.5	.76	.2344	.2997	-.2380	.3000	.7305	20.000
2S-GMM	24	-.5	.76	.3715	.3746	-3310	.3340	-.0120	.0123
3S-EEL	24	-.5	.76	.3608	.3658	-.3225	.3273	-.0116	.0119
3SW-EEL	24	-.5	.76	.3261	.3413	-.2942	.3072	-.0104	.0110
3SWc-EEL	24	-.5	.76	.3288	.3435	-.2960	.3086	-.0106	.0112
CUE	24	-.5	.76	.2533	.3346	-.3125	.3940	-.7314	8.7797

Table 6: Large sample properties in DGP3 ($\rho_1 = .9(1 - \rho_2)$, $\rho_2 = -.65$ and $T = 500$)

				$\gamma_f = .10$		$\gamma_b = .85$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	2.03	.2451	.2572	-.2146	.2251	-.0075	.0079
3S-EEL	8	0	2.03	.1924	.2199	-.1690	.1931	-.0057	.0067
3SW-EEL	8	0	2.03	.1730	.2093	-.1523	.1838	-.0051	.0065
3SWc-EEL	8	0	2.03	.2291	.2445	-.2004	.2138	-.0071	.0077
CUE	8	0	2.03	.1670	.2108	-.1469	.1845	-.0050	.0066
2S-GMM	16	0	2.03	.3066	.3112	-.2678	.2719	-.0095	.0097
3S-EEL	16	0	2.03	.2629	.2768	-.2301	.2422	-.0080	.0085
3SW-EEL	16	0	2.03	.2420	.2630	-.2119	.2302	-.0073	.0081
3SWc-EEL	16	0	2.03	.2650	.2785	-.2316	.2434	-.0082	.0087
CUE	16	0	2.03	.1641	.2300	-.1435	.2013	-.0048	.0072
2S-GMM	24	0	2.03	.3320	.3344	-.2898	.2921	-.0104	.0105
3S-EEL	24	0	2.03	.3083	.3144	-.2696	.2748	-.0095	.0097
3SW-EEL	24	0	2.03	.2976	.3065	-.2602	.2680	-.0092	.0095
3SWc-EEL	24	0	2.03	.3072	.3141	-.2683	.2744	-.0095	.0098
CUE	24	0	2.03	.1847	.2513	-.1629	.2213	-.0055	.0083
2S-GMM	8	.5	1.75	.2421	.2538	-.2121	.2221	-.0074	.0078
3S-EEL	8	.5	1.75	.1962	.2221	-.1723	.1946	-.0059	.0069
3SW-EEL	8	.5	1.75	.1782	.2137	-.1566	.1873	-.0053	.0067
3SWc-EEL	8	.5	1.75	.2292	.2448	-.2005	.2141	-.0071	.0077
CUE	8	.5	1.75	.1741	.2162	-.1530	.1893	-.0052	.0068
2S-GMM	16	.5	1.75	.3010	.3054	-.2631	.2669	-.0093	.0095
3S-EEL	16	.5	1.75	.2664	.2784	-.2332	.2436	-.0081	.0087
3SW-EEL	16	.5	1.75	.2481	.2669	-.2174	.2335	-.0075	.0083
3SWc-EEL	16	.5	1.75	.2681	.2802	-.2344	.2449	-.0082	.0087
CUE	16	.5	1.75	.1779	.2402	-.1563	.2102	-.0053	.0076
2S-GMM	24	.5	1.75	.3251	.3275	-.2841	.2862	-.0101	.0102
3S-EEL	24	.5	1.75	.3097	.3148	-.2707	.2751	-.0095	.0097
3SW-EEL	24	.5	1.75	.3019	.3093	-.2640	.2704	-.0093	.0096
3SWc-EEL	24	.5	1.75	.3092	.3150	-.2701	.2751	-.0095	.0098
CUE	24	.5	1.75	.2089	.2667	-.1844	.2351	-.0063	.0084
2S-GMM	8	-.5	2.36	.2504	.2626	-.2188	.2297	-.0076	.0080
3S-EEL	8	-.5	2.36	.1901	.2192	-.1670	.1925	-.0056	.0067
3SW-EEL	8	-.5	2.36	.1696	.2077	-.1492	.1824	-.0049	.0063
3SWc-EEL	8	-.5	2.36	.2297	.2454	-.2005	.2144	-.0071	.0077
CUE	8	-.5	2.36	.1631	.2093	-.1431	.1833	-.0048	.0065
2S-GMM	16	-.5	2.36	.3113	.2166	-.2718	.2762	-.0097	.0099
3S-EEL	16	-.5	2.36	.2549	.2711	-.2232	.2376	-.0077	.0083
3SW-EEL	16	-.5	2.36	.2317	.2556	-.2032	.2241	-.0070	.0079
3SWc-EEL	16	-.5	2.36	.2608	.2751	-.2278	.2406	-.0080	.0086
CUE	16	-.5	2.36	.1414	.2166	-.1244	.1901	-.0042	.0068
2S-GMM	24	-.5	2.36	.3402	.3429	-.2973	.2998	-.0107	.0108
3S-EEL	24	-.5	2.36	.3058	.3134	-.2678	.2746	-.0094	.0097
3SW-EEL	24	-.5	2.36	.2923	.3035	-.2561	.2660	-.0090	.0094
3SWc-EEL	24	-.5	2.36	.3053	.3138	-.2670	.2746	-.0095	.0098
CUE	24	-.5	2.36	.1604	.2385	-.1417	.2097	-.0048	.0075

Table 7: Finite sample properties in DGP2 ($\rho_1 = .9(1-\rho_2)$, $\rho_2 = -.65/\sqrt{T}$ and $T = 160$)

				$\gamma_f = .850$		$\gamma_b = .100$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	.0004	-.5559	.5772	.0300	.0722	.0526	.0557
3S-EEL	8	0	.0004	-.4828	.5238	.0280	.0739	.0456	.0506
3SW-EEL	8	0	.0004	-.4439	.4985	.0295	.0763	.0415	.0479
3SWc-EEL	8	0	.0004	-.4673	.5095	.0346	.0778	.0431	.0485
CUE	8	0	.0004	-.4487	.5122	.0304	.0779	.0214	.0662
2S-GMM	16	0	.0004	-.6000	.6141	.0368	.0761	.0563	.0588
3S-EEL	16	0	.0004	-.5419	.5693	.0397	.0832	.0501	.0542
3SW-EEL	16	0	.0004	-.4849	.5358	.0446	.0917	.0437	.0506
3SWc-EEL	16	0	.0004	-.4876	.5366	.0481	.0932	.0434	.0503
CUE	16	0	.0004	-.4762	.5544	.0439	.1333	-3.5342	109.09
2S-GMM	24	0	.0004	-.6311	.6417	.0373	.0744	.0594	.0616
3S-EEL	24	0	.0004	-.5865	.6054	.0433	.0830	.0542	.0576
3SW-EEL	24	0	.0004	-.5215	.5621	.0533	.0981	.0462	.0526
3SWc-EEL	24	0	.0004	-.5087	.5476	.0705	.0906	.0329	.0422
CUE	24	0	.0004	-.5019	.5816	.0630	.1922	-3.1929	114.73

Table 8: Large sample properties in DGP2 ($\rho_1 = .9(1-\rho_2)$, $\rho_2 = -.65/\sqrt{T}$ and $T = 500$)

				$\gamma_f = .850$		$\gamma_b = .100$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	.0001	-.5566	.5815	.0097	.0432	.0570	.0600
3S-EEL	8	0	.0001	-.4605	.5099	.0071	.0443	.0473	.0528
3SW-EEL	8	0	.0001	-.4202	.4831	.0070	.0454	.0432	.0499
3SWc-EEL	8	0	.0001	-.4539	.4992	.0095	.0446	.0463	.0514
CUE	8	0	.0001	-.4388	.5061	.0085	.0459	.0449	.0522
2S-GMM	16	0	.0001	-.6005	.6168	.0139	.0452	.0609	.0630
3S-EEL	16	0	.0001	-.5005	.5386	.0114	.0468	.0508	.0552
3SW-EEL	16	0	.0001	-.4377	.4967	.0120	.0483	.0442	.0508
3SWc-EEL	16	0	.0001	-.4436	.4975	.0142	.0484	.0445	.0507
CUE	16	0	.0001	-.3904	.4875	.0114	.0509	.0394	.0502
2S-GMM	24	0	.0001	-.6284	.6411	.0159	.0465	.0636	.0654
3S-EEL	24	0	.0001	-.5403	.5682	.0151	.0492	.0546	.0580
3SW-EEL	24	0	.0001	-.4771	.5217	.0169	.0510	.0478	.0531
3SWc-EEL	24	0	.0001	-.4781	.5206	.0187	.0514	.0476	.0528
CUE	24	0	.0001	-.4121	.5109	.0147	.0540	.0413	.0524

Table 9: Large sample properties in DGP2 (asymmetric error terms)
($\rho_1 = .9(1-\rho_2)$, $\rho_2 = -.65$ and $T = 160$)

				$\gamma_f = .850$		$\gamma_b = .100$		$\lambda = .015$	
	K	ρ	cp	Bias	RMSE	Bias	RMSE	Bias	RMSE
2S-GMM	8	0	15.74	-.2364	.2803	.0260	.0718	.0043	.0073
3S-EEL	8	0	15.74	-.1807	.2367	.0238	.0735	.0030	.0066
3SW-EEL	8	0	15.74	-.1719	.2330	.0240	.0741	.0028	.0065
3SWc-EEL	8	0	15.74	-.1817	.2344	.0290	.0757	.0025	.0062
CUE	8	0	15.74	-.1382	.2110	.0255	.0817	-.0014	.2930
2S-GMM	16	0	15.74	-.3425	.3749	.0345	.0078	.0068	.0093
3S-EEL	16	0	15.74	-.2674	.3137	.0429	.0087	.0043	.0078
3SW-EEL	16	0	15.74	-.2500	.3018	.0440	.0088	.0038	.0076
3SWc-EEL	16	0	15.74	-.2521	.3008	.0462	.0089	.0036	.0075
CUE	16	0	15.74	-.1801	.2820	.0568	.1483	-1.2616	54.238
2S-GMM	24	0	15.74	-.4238	.4378	.0365	.0789	.0089	.0111
3S-EEL	24	0	15.74	-.3543	.3882	.0503	.0905	.0064	.0095
3SW-EEL	24	0	15.74	-.3403	.3787	.0517	.0923	.0059	.0094
3SWc-EEL	24	0	15.74	-.3395	.3771	.0527	.0927	.0058	.0093
CUE	24	0	15.74	-.2749	.3934	.0956	.2366	-10.946	260.22

Table 10: Estimation of the reduced-form of the Hybrid NKPC (GG instrument set)

Estimator	λ	γ_f	γ_b	J_T-stat	IPS_T-stat	K-stat
2S-GMM	0.1338	0.5981	0.3703	13.6726		0.1611
	(0.0483)	(0.0350)	(0.0328)			
	[0.0063]	[0.0000]	[0.0000]	[0.9127]		[0.6882]
CUE	0.1711	0.6544	0.2971	13.6710		0.4329
	(0.0485)	(0.0336)	(0.0325)			
	[0.0006]	[0.0000]	[0.0000]	[0.9127]		[0.5106]
3S-EEL	0.1171	0.6117	0.3615	0.9512	55.2955	0.3027
	(0.0417)	(0.0299)	(0.0290)			
	[0.0057]	[0.0000]	[0.0000]	[0.9512]	[0.0001]	[0.5822]
3Sc-EEL	0.1108	0.6119	0.3627	0.9517	54.9047	0.3016
	(0.0413)	(0.0297)	(0.0288)			
	[0.0082]	[0.0000]	[0.0000]	[0.9517]	[0.0001]	[0.5829]
3SW-EEL	0.1220	0.6142	0.3551	0.9516	55.0016	0.2438
	(0.0422)	(0.0298)	(0.0290)			
	[0.0045]	[0.0000]	[0.0000]	[0.9516]	[0.0001]	[0.6215]
3SWc-EEL	0.1157	0.6144	0.3564	0.9519	54.6957	0.2422
	(0.0418)	(0.0296)	(0.0288)			
	[0.0064]	[0.0000]	[0.0000]	[0.9519]	[0.0001]	[0.6226]

Note: Standard errors appear in parentheses and p-values are in brackets for the null hypothesis that the estimate is equal to zero. The values in the J_T-stat (respectively J_T-stat) column are respectively the test-statistics and the corresponding p-value (in brackets). The last column describes respectively the K-test statistics and their p-values (in brackets).The instrument set GG includes four lags each of inflation, the labor income share, the output gap, the long-short interest rate spread, wage inflation, and commodity price inflation.

Table 11: Estimation of the reduced-form of the Hybrid NKPC (GGLS instrument set)

Estimator	λ	γ_f	γ_b	J_T-stat	IPS_T-stat	K-stat
2S-GMM	0.1267	0.6109	0.3721	5.5093		2.7838
	(0.0658)	(0.0573)	(0.0566)			
	[0.0560]	[0.0000]	[0.0000]	[0.7020]		[0.0952]
CUE	0.1260	0.6203	0.3625	5.7576		0.3097
	(0.0620)	(0.0554)	(0.0543)			
	[0.0441]	[0.0000]	[0.0000]	[0.6744]		[0.5778]
3S-EEL	0.1170	0.6120	0.3695	0.6548	7.6417	2.0212
	(0.0613)	(0.0469)	(0.0479)			
	[0.0583]	[0.0000]	[0.0000]	[0.6548]	[0.4692]	[0.1551]
3Sc-EEL	0.1077	0.6089	0.3749	0.6541	7.6523	2.0258
	(0.0607)	(0.0465)	(0.0474)			
	[0.0781]	[0.0000]	[0.0000]	[0.6541]	[0.4682]	[0.1546]
3SW-EEL	0.1250	0.6165	0.3633	0.6525	7.6751	2.1490
	(0.0619)	(0.0475)	(0.0485)			
	[0.0454]	[0.0000]	[0.0000]	[0.6525]	[0.4658]	[0.1427]
3SWc-EEL	0.1151	0.6132	0.3691	0.6557	7.6293	2.1586
	(0.0612)	(0.0470)	(0.0480)			
	[0.0621]	[0.0000]	[0.0000]	[0.6557]	[0.4705]	[0.1418]

Note: Standard errors appear in parentheses and p-values are in brackets for the null hypothesis that the estimate is equal to zero. The values in the J_T-stat (respectively J_T-stat) column are respectively the test-statistics and the corresponding p-value (in brackets). The last column describes respectively the K-test statistics and their p-values (in brackets).The instrument set GGLS includes four lags each of inflation, and two lags each of the output gap, the labor income share, and wage inflation.

Figure 1: Implied probabilities evaluated at the 3S-EEL estimator