Economics Letters 124 (2014) 341–347

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Estimating aggregate autoregressive processes when only macro data are available

Eric Jondeau[a,*], Florian Pelgrin[b]

[a] Swiss Finance Institute and Faculty of Business and Economics (HEC Lausanne), University of Lausanne, CH 1015 Lausanne, Switzerland
[b] EDHEC Business School, France

HIGHLIGHTS

- The aggregation of individual AR(1) models is an infinite AR process.
- We estimate the aggregate process when only macro data is available.
- A parametric and a minimum distance estimator for the aggregate dynamics are proposed.
- The estimators recover the moments of the distribution of the AR parameters.
- The estimators perform very well, even with finite samples.

ARTICLE INFO

Article history:
Received 22 March 2014
Received in revised form
11 June 2014
Accepted 13 June 2014
Available online 20 June 2014

JEL classification:
C2
C13

Keywords:
Autoregressive process
Aggregation
Heterogeneity

ABSTRACT

The aggregation of individual random AR(1) models generally leads to an AR(∞) process. We provide two consistent estimators of aggregate dynamics based on either a parametric regression or a minimum distance approach for use when only macro data are available. Notably, both estimators allow us to recover some moments of the cross-sectional distribution of the autoregressive parameter. Both estimators perform very well in our Monte-Carlo experiment, even with finite samples.

1. Introduction

Aggregation is a critical and widely acknowledged issue in theoretical and empirical economics research. As noted by Pesaran and Chudik (2014), among the different aspects of the aggregation problem, the identification and estimation of certain distributional features of the micro-parameters from aggregate relations are important issues, especially when only macro data are available (Robinson, 1978; Granger, 1980; Forni and Lippi, 1997). Notably, identifying such features requires the researchers to derive the optimal aggregate function and to make explicit the back out between "macro" and "micro" parameters. Yet, only a few papers have examined the reliability of macro information in circumventing the aggregation bias in the presence of unobserved micro heterogeneity (Lewbel, 1994; Pesaran, 2003; Carvalho and Dam, 2010; Mayoral, 2013). Our paper contributes to this stream of the literature by providing a solution to this problem for autoregressive models when the time-series and cross-sectional dimensions are both large.

We propose two consistent estimation techniques that rely on a flexible parametric specification of the distribution of the micro-parameters and on the estimation of the hyper-parameters of this cross-sectional distribution. The first method is based on maximum likelihood estimation, while the second method is based on minimum distance estimation. Both methods explicitly account for the set of non-linear restrictions that drive the aggregate parameters and allow us to recover reliable information on the distribution of the micro-parameters. Using Monte Carlo simulation, we show that both methods perform very well, even with relatively small samples.

* Corresponding author. Tel.: +41 21 692 33 49.
E-mail addresses: eric.jondeau@unil.ch (E. Jondeau), florian.pelgrin@edhec.edu (F. Pelgrin).

http://dx.doi.org/10.1016/j.econlet.2014.06.012
0165-1765/

2. The model

Consider the random AR(1) model for $i = 1, \ldots, N$:

$$x_{i,t} = \rho_i x_{i,t-1} + v_{i,t}, \quad (1)$$

where ρ_i denotes an individual-specific random parameter and $v_{i,t}$ is an error term. For instance, such dynamics may represent consumption expenditures across households (Lewbel, 1994), consumer price inflation across subindices (Altissimo et al., 2009), real exchange rates across sectors (Imbs et al., 2005), or real marginal cost across industries (Imbs et al., 2011). Innovation $v_{i,t}$ is decomposed into a common component (ϵ_t) and an idiosyncratic (individual-specific) component ($\eta_{i,t}$):

$$v_{i,t} = \kappa_i \epsilon_t + \eta_{i,t}, \quad (2)$$

where κ_i denotes a scaling parameter. The macro variable results from the aggregation of micro-units, with the use of time-invariant nonrandom weights $W_N = (w_1, \ldots, w_N)'$, with $\sum_{i=1}^N w_i = 1$, so that $X_{N,t} = \sum_{i=1}^N w_i x_{i,t}$. The cross-sectional moments of ρ are $\tilde{E}_N(\rho^s) = \sum_{i=1}^N w_i \rho_i^s$, for all $s = 1, 2, \ldots$. Moreover, the following assumptions hold:

Assumption 1. $|\rho| \le c < 1$ almost surely for some constant c. Random parameters have finite variance and higher moments.

Assumption 2. ϵ_t and $\eta_{i,t}$ are white noise processes with mean zero and variance σ_ϵ^2 and σ_η^2, respectively; ϵ_t and $\eta_{i,t}$ are mutually orthogonal at any lag and lead; $\{\epsilon, \eta_i\}$ and $\{\rho_i, \kappa_i\}$ are mutually independent for all i; ρ_i and κ_i are mutually independent; $E(\kappa) = 1$.

Assumption 3. As $N \to \infty$, $\|W_N\| = O\left(N^{-1/2}\right)$ and $w_i / \|W_N\| = O\left(N^{-1/2}\right)$ for all $i \in \mathbb{N}$.

Assumption 1 guarantees that there are no individual unit root parameters that would dominate at the aggregate level (Zaffaroni, 2004). This assumption implies that the limit aggregate (as $N \to \infty$) has a short memory with an exponentially decaying autocorrelation function.[1] Eqs. (1) and (2) together with Assumption 2 provide a parsimonious form of (statistical) cross-sectional dependence, which is common in the aggregation literature (Forni and Lippi, 1997; Zaffaroni, 2004). The aggregation mechanism depends solely on the characteristics of the common component of the error term, i.e., our specification and assumptions rule out the presence of an idiosyncratic component at the aggregate level.[2] Assumption 3 is a granularity condition, which insures that the weights are not dominated by a few of the cross-sectional units (Gabaix, 2011; Pesaran and Chudik, 2014).[3]

3. Aggregate dynamics

Using the moving average (MA) representation in Eqs. (1)–(2), we can straightforwardly show that the aggregate process, $X_{N,t}$, has the following dynamics[4]:

$$X_{N,t} = \sum_{k=0}^{\infty} \left(\sum_{i=1}^{N} w_i \rho_i^k \kappa_i \right) \epsilon_{t-k} + \sum_{k=0}^{\infty} \left(\sum_{i=1}^{N} w_i \rho_i^k \eta_{i,t-k} \right). \quad (3)$$

When N becomes large, by virtue of the strong law of large moments, the limit aggregate dynamics is obtained.

Proposition 1. *Suppose that Assumptions 1–3 hold. Given the disaggregate model defined in Eqs. (1)–(2), the limit aggregate process as $N \to \infty$ has the following dynamics:*

$$X_t = \sum_{s=0}^{\infty} \gamma_s \epsilon_{t-s} \quad (\textit{MA form}), \quad (4)$$

$$X_t = \sum_{s=1}^{\infty} C_s X_{t-s} + \epsilon_t \quad (\textit{AR form}), \quad (5)$$

where $X_{N,t} \xrightarrow{L^2} X_t$ and $\tilde{E}_N(\rho^s) \xrightarrow{\text{a.s.}} \tilde{E}(\rho^s)$ as $N \to \infty$. Parameters γ_s are defined as $\gamma_s = \tilde{E}(\rho^s)$, with $\sum_{s=0}^\infty |\gamma_s| < \infty$. Parameters C_s are defined by $C_0 = 1$, $C_s = \tilde{E}(c_s)$, $\forall s \ge 1$ with $c_1 = \rho$ and $c_{s+1} = (c_s - C_s)\rho$, with $\sum_{s=0}^\infty |C_s| < \infty$.

Proof. Gonçalves and Gouriéroux (1988) and Lewbel (1994). See Appendix A.

The absence of an idiosyncratic component in Eqs. (4) and (5) is a direct consequence of Assumptions 1–3. Eq. (4) shows that the impulse-response coefficients γ_s are the noncentral moments of the random parameter ρ.[5] Eq. (5) shows that aggregation leads to an infinite autoregressive model for X_t (see Robinson, 1978 and Lewbel, 1994). The autoregressive parameters C_s are nonlinear transformations of the noncentral moments of ρ and satisfy the following nonhomogeneous difference equations (for $s \ge 1$), which turn out to be useful in the estimation with only macro data:

$$C_{s+1} = \tilde{E}(c_{s+1}) = \tilde{E}\left(\rho^{s+1}\right) - \sum_{r=1}^{s} C_r \tilde{E}\left(\rho^{s-r+1}\right), \quad (6)$$

and $\sum_{s=1}^\infty C_s = \sum_{s=1}^\infty \tilde{E}(\rho^s)/(1 + \sum_{s=1}^\infty \tilde{E}(\rho^s)) < 1$ almost surely. In addition, the long-run multiplier is given by $1/(1 - \sum_{s=1}^\infty C_s) = \sum_{s=0}^\infty \tilde{E}(\rho^s)$. With the exception of a degenerate distribution for ρ (Dirac distribution), the aggregate dynamics is richer than the individual dynamics because of the nonergodicity of the random AR(1) process. Conversely, when parameters C_s are known or estimated, the cross-sectional moments can be easily deduced. For instance, the cross-sectional mean and variance are $\tilde{E}(\rho) = C_1$ and $\tilde{V}(\rho) = C_2$, respectively, and the standardized skewness and kurtosis are $\tilde{S}(\rho) = (C_3 - C_1 C_2)/(C_2)^{3/2}$ and $\tilde{K}(\rho) = (C_4 - 2C_1C_3 + C_1^2 C_2 + C_2^2)/(C_2)^2$, respectively.

4. Estimation

The estimation approach that was originally proposed by Lewbel (1994) consists in truncating the infinite sums in Proposition 1 and estimating the resulting dynamics:

$$X_{N,t} = \sum_{s=1}^{K} C_s X_{N,t-s} + V_{N,t}, \quad (7)$$

[1] Assumption 1 can be relaxed to allow for long-memory effects. This point is further discussed in Section 5.

[2] The contribution of idiosyncratic shocks through network effects or nongranularity has been discussed in recent papers (e.g., Gabaix, 2011 and Acemoglu et al., 2012).

[3] Our results extend to the case of (time-varying) stochastic weights. Such an extension requires at least that the weights be distributed independently from the stochastic process defining the random variable.

[4] Put differently, it is an ARMA(N, $N-1$) in the absence of common roots in the individual processes (Granger and Morris, 1976).

[5] Noncentral moments $\gamma_s = \tilde{E}(\rho^s)$ of any (nondegenerate) random variable ρ, defined on [0, 1), satisfy: $1 > \gamma_1 \ge \cdots \ge \gamma_s \ge 0$, $\forall s \ge 1$, and $\gamma_s \to 0$ as $s \to \infty$. See Appendix B for additional properties of noncentral moments.