# Firms' Dynamics and Business Cycle: New Disaggregated Data


Lorenza Rossi and Emilio Zanetti Chini*

*University of Pavia*
*Department in Economics and Management*
*Via San Felice 5 - 27100 Pavia (ITALY)*
*e-mail: emilio.zanettichini@unipv.it*



FIRST VERSION: May 2016
THIS VERSION: August 2017



**Abstract**

We provide stylized facts on firms dynamics by disaggregating U.S. yearly data from 1977 to 2013. To this aim, we use an unobserved component-based method, encompassing several classical regression-based techniques currently in use. Our new time series of entry and exit of firms at establishment level are feasible proxies of business cycle. Exit is a leading and countercyclical indicator, while entry is lagging and procyclical. A structural econometric analysis supports the findings of the most recent theoretical literature on firms dynamics.



**JEL Codes:** C13, C32, C40, E30, E32.

**Keywords:** Bayesian VAR, Business Cycle, Disaggregation, Entry and Exit, State Space.



*Corresponding Author.

# 1 Introduction

Firms are one of the pillars of the economy. Their dynamics is an important indicator of the status of the economic activity. As a consequence, having a precise view of the entry and exit of firms from the market is fundamental in order to detect the business cycle phase.

In this respect any investigator could think that more than one dataset is built-up to conduct empirical analysis on this topic. Instead, this is only partially true. Namely, for the US economy, the dataset currently published by Census Bureau is yearly and constitutes the only long-span dataset for empirical studies on the argument; on the other side, the quarterly time series on firm's dynamics published by the Bureau of Labor Statistics (BLS, henceforth), start in 1993:Q2. Hence, the only time series nowadays available for the estimation of macroeconomic models faces with limited – and limiting – frequency. In facts, despite its rich variety of indicators at different level of aggregation, the yearly frequency of Census data limits the number of observations to use when estimating econometric models for business cycle; on the other hand, the starting date in Nineties of BLS series do not allow a comprehensive long-span analysis as required by modern macroeconometric literature. Such an impressive scarcity of macroeconomic data at higher frequency motivates this paper.

We rely on the literature on DSGE models for firms dynamics and its effects on business cycle. At the current state, the theoretical contributions on firms' dynamics are mainly focused on firms's entry[1]. Rossi (2015) and Hamano and Zanetti (2017) cover this theoretical gap in the literature by showing that firms' exit represents an even stronger propagation mechanism of the business cycle than that of firms' entry. Unfortunately, these efforts to give a theoretical explanation of business dynamics face with the above mentioned data limitations.

In this respect, our contribution is twofold. First, we improve the availability and quality of the US data by disaggregating the yearly series provided by the Census Bureau. Second, we use these new data to perform an econometric analysis of the

US business cycle. In particular, we evaluate the performance of our series either via univariate diagnostics and via comparison of our data with the analogue ones by BLS[2].

In the first part of this paper, we consider two different types disaggregation techniques: (i) the one proposed by Chow and Lin (1971), which is still widely used by Statistical Institutes; and (ii) two models based on unobserved components methods (UCM), originally proposed by Proietti (2006). The first technique is a simple OLS estimation of an autoregression with AR(1) errors. The two models based on UCM are instead represented by a unifying state space system and their estimation rely on an Augmented Kalman Filter (AKF), hence allowing generality and flexibility besides maintaining statistical robustness.

We anticipate that the series disaggregated by UCM are considerably more accurate and credible of the ones resulting by applying naive Chow-Lin model. Furthermore, the selecting procedure of the two UC models suggests that series of EXIT should be fitted by an AutoRegressive model of order 1 – AR(1), henceforth – while the series of ENTRY is better represented by an UC representation of an Autoregressive Distributed Lag model of same oder – ADL(1), henceforth. Both the UC models rely on univariate estimations and thus consider a single regressor variable, here represented by the industrial production. Anyway, the dynamics of a single indicator might influence the final estimates of disaggregated process more than expected. In order to circumvent this problem, we re-apply the new (univariate) UCM on all the indicators of MD-FRED dataset and then average the resulting univariate UC series. In this way, we obtain new series for ENTRY and EXIT which are qualitatively very similar to the ones derived from the single indicator, by taking in account of all the information on the US economy[3].

After a comprehensive discussion of data and methods, we use these new disaggregated series for business cycle analysis. We extract the trend and cycle components via standard first difference operator, the Hodrick and Prescott (1997) (HP) and

Baxter and King (1999) (BK) filters. Once the series have been filtered, we classify them as leading or lagging indicators of the business cycle by looking at the maximum absolute value of cross-correlations between the cycle of disaggregated ENTRY and EXIT and that of the real GDP. We will show that ENTRY is a lag and pro-cyclical indicator of the business cycle. On the contrary, the EXIT series is a countercyclical leading indicator.

Finally, we use bayesian techniques to run a structural analysis, and show the responses of the three different proxies of firms creations and firms destructions (that is, the two disaggregated series and the BLS data) – jointly with real GDP and inflation – to a productivity shock. We find that, for all the models considered, firms creation is pro-cyclical and persistent, while firms destruction is countercyclical and overshooting its long-run level for all proxies used[4]. Such responses are then followed by a positive response of the real GDP and also by a decline in the inflation in all the samples considered. These results are also compatible with recent macroeconomic literature on endogenous firms dynamics[5]. The same econometric analysis is repeated using quarterly BLS series, since they are the benchmark of this strand of literature. Finally, different robustness checks confirm our results.

The rest of this paper is organized as follows. Section 2 deals with the problem of data availability on firms dynamics. Section 3 describes the methodology. Then, Section 4 applies the disaggregating techniques on US data on establishments ENTRY and EXIT and investigates the business cycle properties of the disaggregated series. Section 5 discusses the relevance of their business cycle movements for Macroeconomics by estimating a set of Structural (Bayesian) VAR models. Finally, Section 6 concludes. The associated Supplement complements the results of the paper by providing mathematical details on statistical methods here adopted and further results.

# 2 Data

The only official long-span dataset for U.S. is the Business Dynamics Statistics (BDS), published by Census Bureau Research Data Centers. It gives information about total number of firms, establishments and workers, establishments opened, establishments closed, job creation and job destruction and other derived measures at yearly frequency from 1977 up to 2014 (current release). In turn, BDS data summarize the confidential data of Longitudinal Business Databases, a census of business establishments and firms covering all industries in all US[6].

For higher frequency data, two are the main sources available for applied analysts:

1. BUSINESS EMPLOYMENT DYNAMICS from Bureau of Labor Statistics (BLS, henceforth). This source provides a quarterly census of the labor force in private establishments from 1992:Q3 and measures the net change in employment at the establishment level. According to the BLS definition, a net increase (decrease) in employment comes from opening (closing) establishments. Openings are either establishments with positive third month employment for the first time in the current quarter, with no links to the prior quarter, or with positive third month employment in the current quarter following zero employment in the previous quarter; closings are either establishments with positive third month employment in the previous quarter, with no positive employment reported in the current quarter, or with positive third month employment in the previous quarter followed by zero employment in the current quarter. In the course of this paper we will refer to these series with the "OPENINGS" and "CLOSINGS" labels.

   Alternatively, it is also possible to use the two series of establishment birth and death for total private sector, (BIRTHS" and "DEATH" - or B and D - respectively, henceforth) available from the same source from 1993:Q2. These series are direct observations of the number of firms/establishments.

2. ECONOMAGIC. It provides a monthly series on the number of new business incorporations from 1959:M1 to 1996:M9.

Despite the similar nature of the data, and their temporal contiguity, these two series measure different objects. Incorporations concern firm's level of aggregation, while, on the other hand, establishments are often partitions of the firm. Consequently, the observations have different order of magnitude: if one aggregates the Economagic' s monthly series, the resulting values are, approximately, a half of the observed BLS ones. Hence, no interpolation is possible between the BLS and Economagic series. This implies that the above-mentioned macroeconomic literature is forced to stay to the evidence of short samples, or to make theoretical assumptions on the low of motion of firms' entry without any empirical counterfactual. Perhaps more astonishing, this series is the basis of a recent strand of literature[7] despite the fact that the exact definition of incorporation used is not available.

*Ergo*, firms' dynamics is not observed and commonly measured by a proxy - the employment level - quite imperfect. This is evident from a look at the cross-correlation function of these two series with the lags of series of real gross domestic product – RGDP, henceforth[8] – shown in Figure 1: the estimated bars are largely below the critical values of $\pm 0.20$, corresponding to the blue bars. Our contribution is precisely in the production of new time series on firm's dynamics capable to overcome this measurement error problem.

Finally, as the next Section exposes, the disaggregation of a low-frequency time series is strictly related to the choice of an high frequency series to use as benchmark. What time series should be adopted? In principle, any possible macroeconomic time series generically related to the estimation of GDP aggregate. Hence, we adopt the FRED-MD, a new macroeconomic database of 134 indicators of all sectors of the U.S. economy, recently published by Federal Reserve Bank of St. Louis and extensively described in McCracken and Ng (2016)[9]. Namely, we regress the yearly ENTRY (EXIT) series on all these variables in quarterly frequency by in order to obtain a new quarterly series of the regressand. Anyway, as the next sections will document, this solution conveys a time series so similar to the one obtained via univariate regression that further methodological investigation seems us purely academic.

## 3 Temporal Disaggregation

In a temporal disaggregation exercise, measurements of some variable are available only over $s$ consecutive periods, where $s = 4$ if moving from yearly to quarterly, 12 from yearly to monthly and so on. In our case, the annual total of a macroeconomic (flow) variable (i.e. ENTRY/EXIT) has to be redistributed across the quarters using related series that are available at higher frequency (*indicators*).

More in details, the problem of temporal disaggregation can be solved via *interpolation* or via *distribution*. The former consists in estimating of the missing values of a *stock* variable at points in time that have been systematically skipped by the observation process. The latter arises when *flow* variables are in the form of linear aggregates, as for the case of observations available only as totals or as averages over $s$ consecutive periods. Since establishments ENTRY and EXIT are flow variables, temporal distribution represents exactly the solution for our disaggregation problem.

This Section presents the mostly used disaggregation methods. First, we briefly describe the Chow and Lin (1971) method. This relies on a simple OLS regression of an AR(1) process with the observed indicators as exogenous regressors, which distribution function is assumed to be known. For this reason we label this the Chow-Lin Naive Method (CL-NM). We then discuss the Proietti (2006) UCM. This consists of two models: (i) the first one is a UC representation of the Chow-Lin regression model (CL-UCM); (ii) the second one is a more general autoregressive distributed lag (ADL-UCM) model. Both of them are estimated by AKF[10]. Finally, in the last part of this Section, we propose to apply the univariate UCM to MD-FRED dataset, where we repeat the same UCM for all the indicators therein contained. The disaggregated ENTRY (EXIT) will be then computed as a simple average of all quarterly series deriving from disaggregation of each single indicator. Thus, we label this Combination Method.

In what follows, $y_t$ denotes a realization of a (univariate) time series observed at $t = 1 \dots, t, \dots, T$; multivariate time series are denoted in bold and matrices in capital.

## 3.1 The CL-NM

Let assume, for easy of explanation, that we are interested in univariate disaggregation, that is to disaggregate a single time series using a single indicator variable. This is a simple linear OLS regression of the observed $T \times 1$-dimentional process $y_t$ on the vector of the same dimensions of the indicator variable $x_t$ multiplied by the $T \times sT$ disaggregation matrix $D$, i.e.

$$\begin{aligned} D_s' y_t = D_s' x_t' \beta + D_s' u_t, \quad u_t \sim N\left(0, D_s V D_s'\right) \\ u_t = \phi u_{t-1} + \epsilon_t, \;\; |\phi| < 1, \;\; \epsilon_t \sim NID(0, \sigma^2), \end{aligned} \tag{1}$$

where $D_s$ is a block-diagonal matrix of form

$$D_s \doteq \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

for the case in which $s = 4$, that is from yearly to quarterly frequency.

In order to make (1) equivalent to the CL-NM, three noticeable assumptions on $x_t$ must be done:

**CL 1.** *$x_t$ is strictly exogenous;*

**CL 2.** *$x_t$ free of measurement errors;*

**CL 3.** *$x_t$ is cointegrated with $y_t$.*

*Remark* 1. The univariate framework here adopted does not affect the generality of these assumptions, since they hold for all N possible regressors, that is with $\boldsymbol{X_t} = [x_{1,t}, \ldots, x_{N,t}]$ in place of $x_t$.

## 3.2 The Unobserved Component Method

Assumptions CL1 – CL3 are quite restrictive. In particular, the contribution by Harvey (1989) stresses that CL2 is limiting in many common empirical applications. Here we discuss the way to overcome them. To this aim we rely on the following state space representation for the general model for temporal disaggregation:

$$\begin{cases} y_t = z'\alpha_t + x_t'\beta, \\ \alpha_t = \alpha_{t-1} + W_t'\beta + H\epsilon_t \\ \alpha_1 = a_1 + W_1\beta + H\epsilon_1 \end{cases} \tag{2}$$

The first equation is named *measurement equation*, while the second one is the *transition equation*. The vectors $x_t$ and the matrices $W_t$ contain exogenous regressors, corresponding to the indicators, that enter respectively the measurement and the transition equations and zero elements corresponding to effects that are absent from one or the other equations. The initial state vector, $\alpha_1$, is expressed as a function of fixed and known effects ($a_1$), random stationary effects ($H_1\epsilon_1$, where the notation stresses that $H_1$ might differ from $H$), and regression effects, $W_1\beta$.

For what follows the following assumptions are invoked:

**H 1.** $Var(\beta) \to 0$.

**H 2.** $Var(\beta)^{-1} \to 0$.

Assumption H1 means $\beta$ is fixed, but unknown. It holds if it is deemed that the transition process governing the states has started at time $t = 1$. Assumption H2 implies that $\beta$ has an improper distribution with mean 0 and arbitrarily large variance matrix. This holds if the process has started in the indefinite past.

Both these assumptions are just for convenience and can easy be weaken[11]. Under this general framework, the model (1), properly arranged, can be assumed as particular case of (2).

**Case 1** (CL-UCM model)**.** When $\alpha_t$ is a scalar, $z = 1$, $T = \phi$ are $H = 1$, the system (2) degenerates into a linear regression model with AR(1) errors:

$$y_t = z'\alpha_t + x_t'\beta, \quad \alpha_t = \phi\alpha_{t-1} + \epsilon_t, \quad \epsilon_t \sim NID(0, \sigma^2) \tag{3}$$

with $\phi < 1$, $\alpha_1 \sim N(0, \sigma^2(1 - \phi^2))$.

*Remark* 2. The CL-NM model assumes full cointegration of non-stationary elements eventually present in $x_t$. In this case the deterministic component is handled via inclusion of regressors as, e.g., $x_t = 1, t, x_{3t}, \ldots, x_{kt}]'$ and re-arrangement of the process as $y_t = \mu_t + \gamma t + \sum_j \beta_j x_j + \alpha_j$ with $\mu$ and $\gamma$ being the first two elements of $\beta$.

**Case 2** (ADL-UCM model)**.** If $z' = 1$, $T = \phi$, $H = 1$, $W = [1, t, x_t', x_{t-1}']$, $\beta = [m, g, \beta_0', \beta_1']'$, the system (2) degenerates into an ADL model:

$$y_t = \phi y_{t-1} + m + gt + x'\beta_0 + x_{t-1}'\beta_1 + \epsilon_t, \tag{4}$$

which initial conditions are, under stationarity assumption: $\alpha = 0$, $W_1 = \frac{1}{1-\phi}[1, \frac{1-2\phi}{1-\phi}, x_1', x_1']$, $H_1 = \frac{1}{\sqrt{1-\phi^2}}$. Still, changes are needed if assuming non-stationarity; see Proietti (2006) for further details.

The more general ADL-UCM can nest also the CL-UCM. Indeed:

*Remark* 3. The ADL-UCM collapses to the CL-UCM if and only if:

$$\beta_1 = \phi\beta_0. \tag{5}$$

Thus, (4) can be rewritten as:

$$y_t = x_t'\beta_0 + \alpha_t, \quad \alpha_t = \phi\alpha_{t-1} + \epsilon_t \tag{6}$$

so that the ADL(1,1) model nests a CL model with stationary AR(1) errors. Similarly, when $\beta_1 = 0$, the model become an ADL(1,0).

### 3.2.1 State Space Representation

The issue of temporal disaggregation arise as a modification of system (2). In order to understand this, notice that (low-frequency) data are a sum of $s$ consecutive values, available at time $t = 1, 2, \ldots, [n/s]$, where $[\cdot]$ denote the integer part of the number $n/s$. Then the following holds:

**Proposition 1.** Let define the cumulator variable as:

$$y_t^c \doteq \psi_t y_{t-1}^c + \psi_{t-1}, \quad \psi_t^c \doteq \begin{cases} 0 \ \text{ if } \ t = s(\tau-1)+1, \ \text{ with } \ \tau = 1, \ldots, [n/s] \\ 1 \ \text{ otherwise} \end{cases} \tag{7}$$

Then the State Space System

$$\begin{cases} y_t = z^{*\prime}\alpha_t^* \\ \alpha_t^* = T_t\alpha_{t-1}^* + W_t^*\beta + H^*\epsilon_t, \\ \alpha_1^* = a_1^* + W_1^*\beta + H_1^*\epsilon_1, \qquad \epsilon_t \sim NID(0,\sigma^2), \quad \beta \sim N(b, \sigma^2 V), \end{cases} \tag{8}$$

where:

$$z^* = \begin{bmatrix} 0' \\ 1 \end{bmatrix}, \ T_t^* = \begin{bmatrix} T & 0 \\ z'T & \psi \end{bmatrix}, \ W_t^* = \begin{bmatrix} W_t \\ z'W_t + x_t' \end{bmatrix}, \ H_t^* = \begin{bmatrix} H_t \\ z'H_t + x_t' \end{bmatrix}, \tag{9}$$

$$\alpha_1^* = \begin{bmatrix} a_1 \\ z'a_1 \end{bmatrix}, \ W_1^* = \begin{bmatrix} W_1 \\ z'W_1 + x_1' \end{bmatrix}, \ H_1^* = \begin{bmatrix} H_t \\ z'H_1 + x_1' \end{bmatrix} \tag{10}$$

converts the disaggregation into a problem of estimation of a latent component model with missing observation. This is the system which is estimated for disaggregation of ENTRY and EXIT series, as shown in next Section 4.

*Proof.* This occurs simply replacing $y_t = z\alpha_t$ in (7), substituting the transition equation and re-writing:

$$y_t^c = \psi_t y_{t-1}^c + \psi_{t-1} + z' T\alpha_{t-1} + (z' W_t + x_t\beta) + z' H\eta_t, \tag{11}$$

and finally, substituting this expression in the state vector, with $\alpha_t^* = [\alpha_t', y_t^c]'$ $\square$

#### 3.2.2 Estimation

Estimation of system (8) is done by Maximum Likelihood. Depending on the presence of diffuse elements in matrix $W$ or not, two estimators are possible:

**Fixed**: the ML estimators of $\beta$ and $\sigma^2$ are:

$$\hat{\beta} = -S_{n+1}^{-1} s_{n+1} \quad , Var(\hat{\beta}) = S_{n+1}^{-1}, \quad \hat{\sigma}^2 = \frac{q_{n+1} - s_{n+1}' S_{n+1}^{-1} s_{n+1}}{[n/s]}, \tag{12}$$

with profile Log-Likelihood:

$$\mathcal{L}_{\mathcal{F}} = -0.5[d_{n+1} + [n/s](\ln \hat{\sigma}^2) + (\ln \hat{\sigma}^2 + \ln 2\pi + 1)]. \tag{13}$$

**Diffuse**: $\hat{\beta}$ and $\hat{\sigma}^2$ unmodified and $\hat{\sigma}^2 = \frac{q_{n+1} - s_{n+1}' S_{n+1}^{-1} s_{n+1}}{[n/s]-k}$ and the diffuse profile Log-Likelihood:

$$\mathcal{L}_{\mathcal{D}} := -0.5[d_{n+1} + [n/s - k](\ln \hat{\sigma}^2) + (\ln \hat{\sigma}^2 + \ln 2\pi + 1) + \ln |S_{n+1}|] \tag{14}$$

In both of them, the parameters $\beta$ can be concentrated out of the likelihood function, whereas the diffuse case is accommodated by simple modification of the likelihood.

*Remark* 4. $S$, $s$, $q$ are outcomes of the augmented Kalman Filter (KF). This algorithm, introduced by de Jong (1991), enables exact inferences in the presence of fixed and diffuse regression effects. In order to make it operational in our system (2), the usual KF equations are augmented by additional recursions which apply the same univariate KF to $k$ series of zero values, with different regression effects in the state equation, provided by $W_t$. The outputs vectors and matrices of the augmented KF are the basis for the de Jong (1989) augmented Smoothing algorithm. This refers to the estimation of the state vector $\alpha_t$ and the disturbance vector $u_t$ using information in the whole sample rather than just past data. Smoothing is an important feature because it is the basis for diagnostic checking for detecting and distinguishing between outliers and structural changes using auxiliary residuals. The annexed Supplement provides mathematical details.

### 3.2.3 Combination

The UCM disaggregation previously described holds for a single time series, that is, In terms of system (8), we are assuming $x_t$ is $[T \times 1]$ vector. Such an assumption is hard to justify in many applications, being the resulting disaggregated time series $y_t$ potentially too much depending from the indicator variable.

In order to mitigate this problem, we consider $\mathbf{X_t} = [\mathbf{x_1}, \ldots, \mathbf{x_n}, \ldots, \mathbf{x_N}]$ a $[T \times N]$ matrix containing the $N$ indicators of MD-FRED dataset and run the UCM for disaggregation machinery to all the $N$ elements of $X_t$. In this way, one can use informations from all the sectors of the economy. Consequently, we produce $N$ differently disaggregated processes collected in $[T \times N]$ matrix $\hat{\mathbf{Y}}_\mathbf{t} = [\hat{\mathbf{y}}_\mathbf{1}, \ldots, \hat{\mathbf{y}}_\mathbf{n}, \ldots, \hat{\mathbf{y}}_\mathbf{N}]$, with $y_n$ representing the $[T \times 1]$ vector of the process disaggregated via system (2) which in turn corresponds to the $n-$th indicator of $X_t$ and "$\hat{}$" denoting the fact that the series is the result of an estimation of (8) and (12)–(13) or (12)–(14)

Then the combination of all disaggregated processes contained in $Y_t$ is a simple average:

$$\hat{y}_t^{\,C} \doteq \frac{1}{N}\sum_{n=1}^{N} y^s_{t,n}, \tag{15}$$

with $y^s_{t,n}$ denoting the time series at time $t$ disaggregated according to the univariate UCM using the $n$-th indicator. The equation (15) conveys the series labeled $AVEntry$ and $AVExit$ – to underline the fact that they are an average of many single processes.

# 4 Business Cycle Analysis

A graphical inspection of the series estimated from all the three models considered in previous section is shown in Figure 2. In particular, we notice that the series obtained by CL-NM is completely different to the others at the end of sample (for example in 2013 the annual average level of EXIT is 170,000 in the CL-NM against 188,000 for the equivalent series measured by the CL-UCM). Another difference between the CL-NM and UCM, emerges when looking at the correlations between our disaggregated series of ENTRY (using different techniques) and the quarterly data of establishments OPENINGs (BIRTH) and, in a similar fashion, between disaggregated series of EXIT and the series of CLOSING (DEATH). These are reported in Table 1: there is a large spread between correlations of ENTRY and EXIT measures (0.23 and 0.70 in mean, respectively); moreover, the correlations of our series obtained CL-NM and the OPENING/CLOSING proxies are 40% lower than the UCM equivalent ones. In addition to this, the Johansen (1991) test rejects the null hypothesis of one cointegrating relation between the series of ENTRY and IP – properly aggregated – strongly[12]. Being full cointegration of indicators a requirement for applying the UCM, this finding ends our investigation on the CL-NM, see Remark 2 in Section 3.2.

The series estimated via CL-UCM specification are nicely smooth. Interestingly, this "smoothness" is lower if ADL specification is selected. This is immediately visible in periods of recession. Let consider the example of ENTRY during the recession phase of 1990: the series measured via CL-UCM passes from 178,000 in 1990:Q4 to 172,000 in 1991:Q2, while the ADL-UCM one ranges from 181,500 to 167,000 – that is, the reduction in entry is more than the double, in absolute value, under recession. The quasi-noisy behavior of ADL specification characterizes all the span of the series. Thus, the difference in the estimates deriving from the two specifications of the UCM is not negligible; see also the graphical comparison in Figure 3.

What specification of UCM should be chosen? This is a typical selection problem that can be solved via simple Likelihood-Ratio (LR) test. According to its high LR-statistic (4.3) reported in Table 2, the model chosen for ENTRY is the ADL-UCM. On the contrary, the LR-statistics of EXIT (1.5) is not able to reject the CL-UCM specification.

The LR-test is a criterion to select the model specification to use when combining the data disaggregated by the same UCM from using all the 134 indicators of FRED-MD. This leads to new time series qualitatively very similar to the ones derived from a single indicator previously analyzed, as confirmed by Figure 4.

We are ready to use our disaggregated series of ENTRY and EXIT for business cycle analysis. As Figure 5 shows, all the series have the same path in both trend and cycle components, and just in the trend some difference is visible between the two filters used; the cycle component of univariate and Combination UCM deliver is substantially coincident. If instead we repeat the same graphical check on BLS data, the evidence gives a different picture, which is represented by Figure 6: the BK filter conveys a cycle considerably smoother than the HP one. On the other hand, the HP-filtered OPENINGS series has a wildly noisy trend and a short cycle (very similar to the one of BIRTHS), while the CLOSINGS and DEATHS are more nicely smooth and have a dynamics different in terms of peaks and troughs; it is interesting to notice that BIRTHS (DEATHS) are more (almost) than the double of OPENINGS (CLOSINGS) in the 2007–2009 recession. These differences in the cycles can be better appreciated if comparing the series from all different sources; this is done in Figure 7: our series anticipate the recession phase in ENTRY, albeit BLS data are "leading" in EXIT measures.

At the light to these findings, we assert that the nature of the BLS proxies is different from our disaggregated series. Hence, we need to definitively classify them as leading or lagging indicators of the economic activity. To this aim, we compute the maximum absolute value cross-correlations between the cycle of disaggregated series and that of the RGDP. It seems reasonable to limit our search for the corresponding maximum absolute cross-correlation to a range between lags and leads of 6 quarters for the ENTRY and 8 quarters for the EXIT[13]. According to Table 3 we find the following results:

- the disaggregated ENTRY series from BDS is generally a lag and pro-cyclical indicator of the business cycle with maximum absolute cross-correlation at the lag 3 for the BK and the HP filter and at the lag 8 for the linear difference filter, which constitutes the only difference with AVEntry series.

- On the contrary, the AVExit is negatively correlated with RGDP and leads the cycle with a maximum cross-correlation at lag 6 for the BK and the HP filter and at lag 7 for the linear difference filter. Moreover, the equivalent Univariate UCM series are considerably different: the BK-cycle is counter-cyclical and leading; instead, the opposite holds for the HP cycle (pro-cyclical with two years of lagging).

- The contemporaneous correlations with RGDP are positive for ENTRY and negative for EXIT. However, they are not statistically significant. Differently, both of the two BLS proxies are generally pro-cyclical. Namely, OPENINGS and BIRTHS are coincident indicators, while CLOSINGS (DEATHS) are generally lagging (leading).

# 5 Structural Analysis

Once the statistical properties of the new disaggregated series have been investigated, it is possible to use them for structural analysis. To this aim, we now estimate small Bayesian VAR (BVAR) models and show the impulse response functions (IRFs - henceforth) to orthogonalized shocks to labor productivity as well as to the TFP (IRFs to TFP shocks are shown in the Supplement).[14] We use data on Real GDP, Inflation, and the three alternative proxies of firms' dynamics, that is:

1. the disaggregated series of firms AVEntry and AVExit obtained with the multivariate UC method;
2. the series of firms ENTRY and EXIT obtained with the univariate UC method, and finally;
3. the BLS quarterly series of firms BIRTHS and DEATHS.

The Bayesian approach to structural analysis has been considered for its versatility and capability to address the issue of short sample size. In fact, it avoids sampling errors in estimation of IRF bands that may occur when sample is short or, equivalently, when the model is highly over parameterized, see Sims and Zha (1998). We estimate five different BVAR specifications. In particular, we consider five different combination of Entry and Exit proxies (labeled M1–M5, henceforth). Table 4 summarizes the model definitions and gives a complete description of the samples. Several important remarks are necessary:

(i) we consider the sample size 1993:Q2-2013:Q4 for the BVAR estimated with the BLS series of BIRTHS and DEATHS, while a longer sample (1983:Q1-2013:Q4) is instead used for the BVAR estimated using the disaggregated (AV)Entry and (AV)Exit. Notice that, we exclude the period of high inflation and the monetary regime change of the early eighties and we consider only the Volcker-Greenspan-Bernanke period.

(iii) Two BVAR using the disaggregated series for the sample 1993Q2-2013Q4 are also considered to compare their analogues which adopt BLS data.

(iv) All the variables under consideration are expressed in logarithms of the levels. Precisely, for each variable $Y$ we adopt $LY$ to label its log-transform[15]. As it is now standard in the literature, this implicitly allows for the possible presence of cointegrating relations, without imposing restrictions on the long-run properties of the model, see Sims *et al.* (1990).

The reduced for of the BVAR is:

$$\boldsymbol{y_t} = \boldsymbol{A_1 y_{t-1}} + \boldsymbol{A_2 y_{t-2}} + ... + \boldsymbol{A_p y_{t-p}} + \boldsymbol{u_t}, \tag{16}$$

where $\boldsymbol{A_j}$, $(j = 1 \dots p)$ are $(N \times N)$ coefficient matrices and $\boldsymbol{u_t} = [u_{LabProd,t,}, u_{RGDP,t}, u_{ENTRY,t}, u_{EXIT,t}]'$ is a white noise vector of time series with $\boldsymbol{u_t} \sim (\boldsymbol{0}, \boldsymbol{\Sigma_u})$.

For each model considered, we estimate a BVAR(2) using Minnesota Priors of 0.8 on the autoregressive coefficient of the first lag for the distribution of parameters. These priors are justified by the short sample size. In fact, in this framework, it is assumed that the VAR residual variance-covariance matrix $\boldsymbol{\Sigma_u}$ is known; here we make use the entire variance-covariance matrix of the VAR system estimated by OLS[16]. Normal Diffuse Priors are also used for robustness check. Finally, we choose a lower triangular Cholesky identification and order LLabProd first, such that on impact shocks to productivity affect the other variables, while shocks to the other variables do not affect productivity on impact. This ordering has been widely used in the literature, see Bloom (2009) *inter alia*. In what follows, we consider responses to 1 percent shock to the labor productivity.

## 5.1 Results

The IRFs resulting from the BVAR under Minnesota Priors are plotted in Figures 8 – 12, while Figures 13 – 17 show the same objects obtained by assuming Normal

Diffuse Priors[17]. The median responses of the endogenous variables to one-standard-deviation increase in the innovations to labor productivity are depicted by solid lines, while dashed lines represent 84 and 16 percent credible intervals. Three findings are noticeable:

1. The real GDP increases in response to a productivity shock, while inflation falls down; this is exactly the picture depicted by New Keynesian theory.

2. Firms creation is pro-cyclical and persistent for all the proxies considered in the five different specifications of the BVAR model. As suggested by recent the theoretical model described in Rossi (2015): (i) firms destruction is countercyclical (for all the proxies considered); and, remarkably, (ii) it overshoots it long run level in the medium run.

3. Our results are robust to the use of the Normal Diffuse Priors. In fact, the IRFs have the same patterns of the ones obtained with Minnesota priors, even though, as expected, the credible interval are larger on impact. This is particularly evident for the disaggregated series of EXIT.

Overall, our BVAR analysis confirms our statistical findings reported in Table 3, hence suggesting that the series of entry and exit are, respectively, procyclical and countercyclical if conditioned to a productivity shock. We underline that the different proxies of firms destructions seems to be less inertial than firms creations and overshoot their long run level in the medium run. Consequently, our disaggregated, considerably longer series are an interesting alternative to the use of the BLS series in the empirical analysis.

The Supplement shows the IRFs of the five models M1-M5, where all series are in log-difference instead of logs. Here, we consider both the IRFs to a one standard deviation labor productivity shock, as well as to a TFP shock. Overall, we can state that our results are robust to the use of the series in log-difference, as well as to the introduction of the alternative measure of productivity.

# 6 Conclusions

Recent advances in macroeconomic theory make the availability of new time series data on firms' dynamics a hot issue. This need is here satisfied by applying an unobserved component-based temporal disaggregation method to data on entry and exit of firms at establishment level.

Our new time series of entry and exit of firms at establishment level are feasible proxies of business cycle. In particular, EXIT is a leading and countercyclical indicator, while ENTRY is lagging and pro-cyclical. Moreover, our new quarterly BDS series on firms' dynamics shows that a TFP shock is associated to a negative and persistent response of EXIT and a positive and persistent response of ENTRY.

The availability of new macroeconomic data allows us to gain additional information to use for our estimates. No matter of this, the similarity of the combination of 134 disaggregated series deriving from the FRED-MD dataset with the ones resulting from univariate UCM poses a doubt on the effectiveness of this last ad-hoc combination strategy. We think that such a result can be explained from the simple fact that many of the variables in the FRED-MD have paths mutually different. This somehow compensates many of the potential differences in series singularly derived from these indicators. Thus, we recommend caution in the use of on the disaggregation methods based on large datasets without having a proper, economically meaningful selection of the indicators. This could be an interesting development in light of new advances in variable selection algorithms as in Bai and Ng (2009).

Finally, we are aware that our disaggregated series could be affected by nonlinearity. This is partially confirmed by Zanetti Chini (2017), and could be due to the fact that we used a nonlinear variable (the industrial production) as indicator. Thus, we are confident that a properly setted nonlinear structure could be useful to implement descriptive properties of the process and to forecasting aims.

# Notes

1 In their seminal article, Bilbiie *et al.* (2012) – BGM, henceforth – introduce a DSGE model with endogenous firms' entry, according to which the sluggish response of the number of producers, due to the sunk entry costs, generates a new, potentially important endogenous propagation mechanism for real business cycle models; see also Bergin and Corsetti (2008); Jaimovich and Floetotto (2008); Etro and Colciago (2010); Colciago and Rossi (2012); Lewis and Poilly (2012); Siemer (2014); Bergin *et al.* (2014); Casares and Poutinau (2014); La Croce and Rossi (2015). These papers consider an exogenous and constant exit probability of firms from the market. Ergo, they are not able to disentangle the role of firms exit with respect to firms' entry.

2 All the series - the disaggregate ones and the equivalent quarterly ones - are here generically labelled as "ENTRY" and "EXIT", while, in the rest of the paper, such labels will be substituted to correctly identify each single series used.

3 We must remark that averaging many different disaggregated processes assumes implicitly that each indicator has equal weight in defining the combined time series. Despite the fact this choice could be not the best, we will adopt this strategy as robustness check. More refined methods that use factor analysis like the contributions by see Frale *et al.* (2011) and Grassi *et al.* (2015) are available and constitute a possible development in this sense.

4 This peculiarity is assumed in the model by Rossi (2015), which is here confirmed by data.

5 See, *inter alia* Bilbiie *et al.* (2012); Lewis (2009); Etro and Colciago (2010); Colciago and Rossi (2015).

6 For more details, see the Census web page, where data are available. When this paper was initialized, we adopted the 2015 release the time span of the data was up to 2013. We maintain this release to avoid problems of variation in the estimates that often characterizes official datasets, as in our case.

7 See Bergin and Corsetti (2008); Lewis and Poilly (2012); Bergin *et al.* (2014); Lewis (2013) *inter alia.*

8 Data source: FRED, Federal Reserve Bank of St. Louis.

9 The FRED-MD is at monthly frequency and, at the time of the settlement of our investigation, it was the only available online. Since in our application the focus is in quarterly data, we aggregated the original data available from the release 2015-04. At the present date, the quarterly version of the dataset, FRED-QD, is ready for the use. All the monthly releases and

the quarterly version can be downloaded at the Michael McCraken's web-site.

10 Both of the models considered rely on a single regressor, or indicator variable, represented by the series of Industrial Production Index, downloaded from FRED of St. Louis.

11 In particular, $\beta_t$ could be assumed to vary over time, see Durbin and Koopman (2012) for more complex specification of the latent component model. Anyway, we find that our data do not makes such a complex dynamics more useful than the simpler one here adopted.

12 The Johansen test is a simple likelihood-ratio test on the hypothesis that matrix $Z = [x : y]$ has rank at least zero; if rejected, it is possible to continue the investigation for higher ranks up to the hypothesis of rank of $n-1$, for $n$ denoting the number of variable in the VAR system; here, we assumed a bivariate VAR so the only two alternative hypothesis coincides with no cointegration vs perfect cointegration. In particular, our application of the test via CATS for RATS software, with 2,500 bootstrap replications, leads to a trace test statistic of 43.766 for the hypothesis of rank 1 and a $p$-value of less than 1%, and consequently it is strongly rejected. See Juselius (2006) for further methodological details.

13 We enlarged the number of possible lags/leads whenever it was not possible to find a maximum inside this range in order to avoid spurious results. In facts, whenever the duration of the business cycles is short, a variable that leads (lags) the reference cycle by several months can be wrongly classified as lagging (leading) since it can be closer to the previous cycle than to the next; see Altissimo *et al.* (2001, 2010) for a discussion of this problem. In our application, Figure 5 makes us able to notice that the cycle of ENTRY is shorter than the cycle of EXIT. *Ergo*, the choice of a different number of leads/lags for the two disaggregated series.

14 Here, we consider labor productivity for two reasons. i) it is the variable considered in the theoretical model by Bilbiie *et al.* (2012); Rossi (2015); Colciago and Rossi (2015); **?**, among others. ii) It is available in levels and in quarterly terms from FRED database, while TFP is available only in annual terms. Nevertheless, for robust check we use the standard TFP measure computed by Basu *et al.* (2006). This series is quarterly, but available only in log-difference terms. For this reason we consider the TFP only in the BVAR estimated in log-deviations. In the Technical Appendix,these estimations will be compared with the same BVAR where also labor productivity is in logarithm difference.

15 For example, in M1 the $N$ vector of observable time series is $\boldsymbol{y_t} = [LLabProd_t, LCPI_t, LRGDP_t, LAVENTRY_t, LAVEXIT_t]'$, where LLabProd is the logarithm of labor productivity, $LCPI_t$ is the logarithm of Core CPI Index, while $LRGDP_t$, the logarithm of real GDP, LAVEntry and LAVExit are the logarithms of the disaggregated series of AVEntry and AVExit (that is

the ones obtained using the multivariate UC method) and the vector length is $T$ with $T = 122$ at the full sample. All these series has been dowloaded from FRED database.

16 Since the model estimates all the equations simultaneously, the assumption of a diagonal matrix of the original Minnesota prior, which in turn implies independence between the VAR coefficients of different equations and justified at that time of limited computational power as it estimates the model equation by equation, is not necessary.

17 Also in this case we set a prior of 0.8 on the autoregressive coefficient of the first lag.

# Acknowledgments

This paper has been supported by the Project "MACFINROBODS" supported by the European Union. The Authors are grateful to Tommaso Proietti for having shared their "TDS" package in Ox; this package has been properly modified by the Authors, who are responsible for any error or imprecision in the calculations. Moreover, the Authors thank Barbara Annicchiarico, Carolina Castagnetti, Efrem Castelnuovo, Cecilia Susanna Jona-Lasinio, Vivien Lewis, Jean Palate, Anthony Savagar, Chris Sims and all participants to ERCIM-CFE2015 in London and 2nd MACFINROBODS in Bruxelles and the Internal Economic Seminar at University of Pavia. Finally, Emilio Zanetti Chini is in debt with the doctors, assistants and nurses of the Department of Emathology of Policlinico "S. Matteo" of Pavia, without whose (free) cares this paper would possibly remain unfinished.

# References

Altissimo, F., Bassanetti, A., Cristaldoro, R., Forni, M., Lippi, M., Reichlin, L. and Veronese, G. (2001). EUROCOIN: A Real Time Coincident Indicator of the Euro Area Business Cycle. CEPR Working Paper No. 3108.

—, Cristaldoro, R., Forni, M., Lippi, M., Reichlin, L. and Veronese, G.

(2010). New EUROCOIN: Tracking Economic Growth in Real Time. *The Review of Economics and Statistics*, **92**, 1024–1034.

Bai, J. and Ng, S. (2009). Boosting diffusion indices. *Journal of Applied Econometrics*, **24**, 607–629.

Basu, S., Fernald, J. and Kimball, M. (2006). Are technology improvements contractionary. *American Economic Review*, pp. 1418–1448.

Baxter, M. and King, R. (1999). Measuring business cycles: approximate band-pass filters for economic time series. *The Review of Economics and Statistics*, **81**, 575–593.

Bergin, P. and Corsetti, G. (2008). The extensive margin and monetary policy. *Journal of Monetary Economics*, **55**, 1222–1237.

—, Feng, L. and Lin, C. (2014). Financial Frictions and Firms Dynamics. NBER Working Paper no 20099.

Bilbiie, F., Ghironi, F. and Melitz, M. (2012). Endogenous Entry, Product Variety, and Business Cycle. *Journal of Political Economy*, **120**, 304–345.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, **77**, 623–685.

Casares, M. and Poutinau, J. (2014). A DSGE Model with Endogenous Entry and Exit. Carleton Economics Paper no. 14-06.

Chow, G. and Lin, A. (1971). Best Linear Unbiased Interpolation, distribution and extrapolation of time series by related series. *The review of Economics and Statistics*, **53**, 372–75.

Colciago, A. and Rossi, L. (2012). Firms Entry, Oligopolistic Competition and Labor Market Dynamics. DNB Working Paper (Bank of Netherland).

— and — (2015). Firms Entry, Oligopolistic Competition and Labor Market Dynamics. DNB Working Paper No. 465.

de Jong, P. (1989). Smoothing and interpolation with the state-space model. *Journal of American Statistical Association*, **84**, 1085–1088.

— (1991). The diffuse Kalman Filter. *Annals of Statistics*, **19**, 1073–1083.

Durbin, J. and Koopman, S. (2012). *Time Series Analysis by State Space Models*. Oxford, UK.

Etro, F. and Colciago, A. (2010). Endogenous Martket Structure and the Business Cycle. *The Economic Journal*, **120**, 1201–1233.

Frale, C., Marcellino, M., Mazzi, G. and Proietti, T. (2011). Euromind: a monthly indicator of euro area economic conditions. *Journal of Royal Statistical Association, ser. A*, **174**, 439–470.

Grassi, S., Proietti, T., Frale, C., Marcellino, M. and Mazzi, G. (2015). EuroMInd-C: Disaggregate Monthly Indicator of Economic Activity for the Euro Area and member countries. *International Journal of Forecasting*, **31**, 712–732.

Hamano, M. and Zanetti, F. (2017). Endogenous Product Turnover and Macroeconomic Dynamics. *Review of Economic Dynamics*, **26**, 263–279.

Harvey, A. (1989). *Forecasting, Structural Time Series Model and the Kalman-Filter*. Cambridge, UK: Cambridge University Press.

Hodrick, R. and Prescott, E. (1997). Postwar us business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, pp. 1–16.

Jaimovich, N. and Floetotto, M. (2008). Firm Dynamics, Mark-up Variations and the Business Cycle. *Journal of Monetary Economics*, **55**, 1238–1252.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, **59**, 1551–1580.

Juselius, K. (2006). *The Cointegrated VAR Model – Methodology and Applications*. Oxford, UK.

La Croce, C. and Rossi, L. (2015). Firms Endogenous Entry and Monopolistic Banking in a DSGE Model. *Macroeconomic Dynamics*, **Forthcoming**.

Lewis, V. (2009). Business Cycle Evidence on Firm Entry. *Macroeconomic Dynamics*, **127**, 324–348.

— (2013). Optimal Monetary Policy and Firm Entry. *Macroeconomic Dynamics*, **17**, 1687–1710.

— and Poilly, C. (2012). Firm entry, markups and the monetary transmission mechanism. *Journal of Monetary Economics*, **59**, 670–685.

McCracken, M. and Ng, S. (2016). FRED-MD: A Monthly Database for Macroeconomic Research. *Journal of Business & Economic Statistics*, **34**, 574–589.

Proietti, T. (2006). Temporal Disaggregation by state space methods: Dynamic regression methods revisited. *Econometric Journal*, **9**, 357–372.

Rossi, L. (2015). Endogenous Firms' Exit, Inefficient Banks and Business Cycle Dynamics. DEM Working Paper no. 99, University of Pavia.

Siemer, M. (2014). Firm Entry and Employment Dynamics in the Great Recession. FEDS Working Paper No. 2014-56.

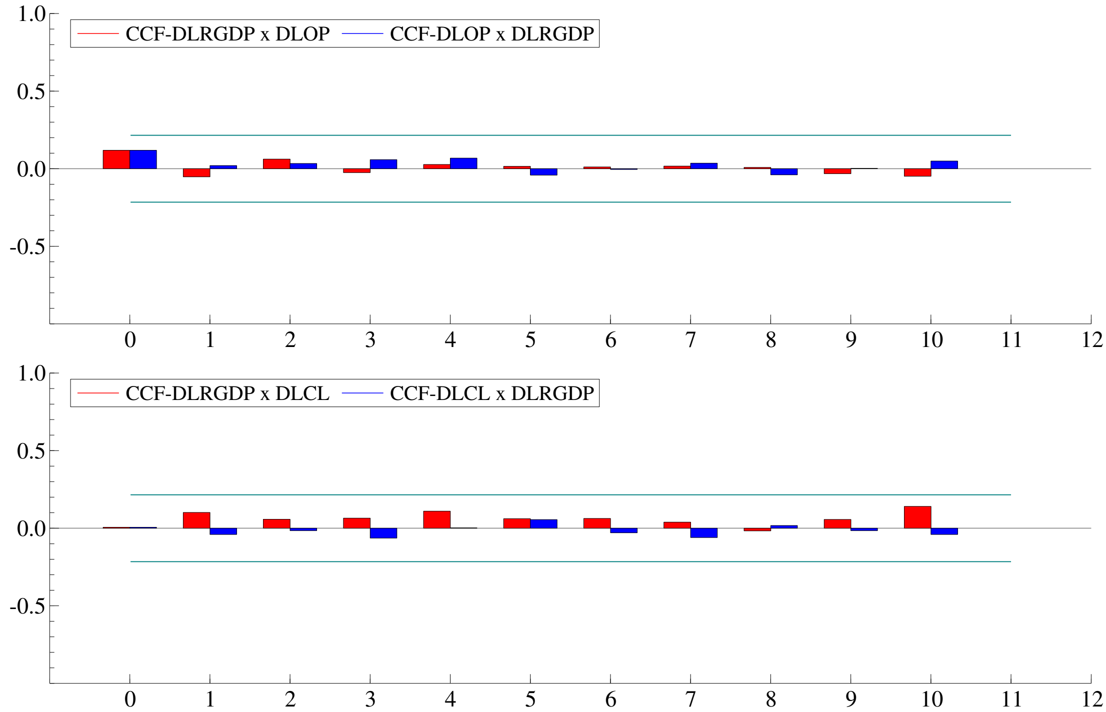
Sims, C., Stock, J. and Watson, M. (1990). Inference in linear time series models with some unit roots. *Econometrica*, **58**, 113–144.

— and Zha, T. (1998). Bayesian Methods for Dynamic Multivariate Models. *International Economic Review*, **39**, 949–968.

Zanetti Chini, E. (2017). Generalizing Smooth Transition Autoregression. DEM Working Paper No. 138, University of Pavia.

# 7 Figures and Tables

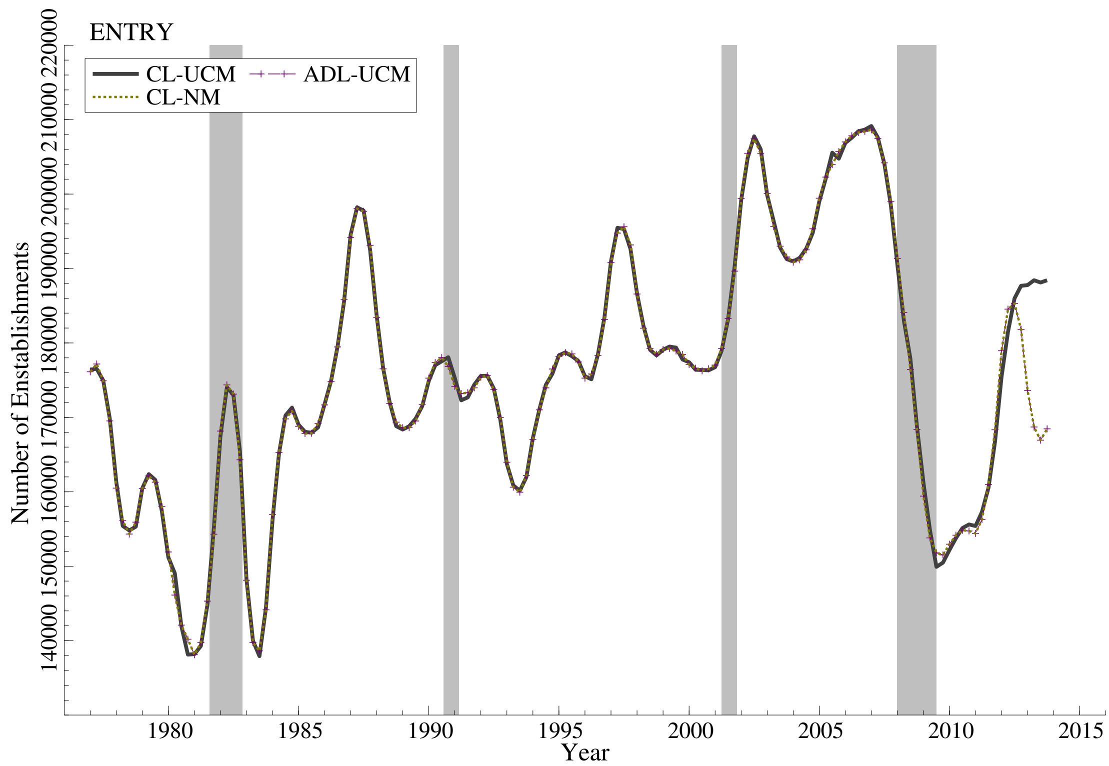
**Figure 1:** Cross-correlation function of BLS data with RGDP



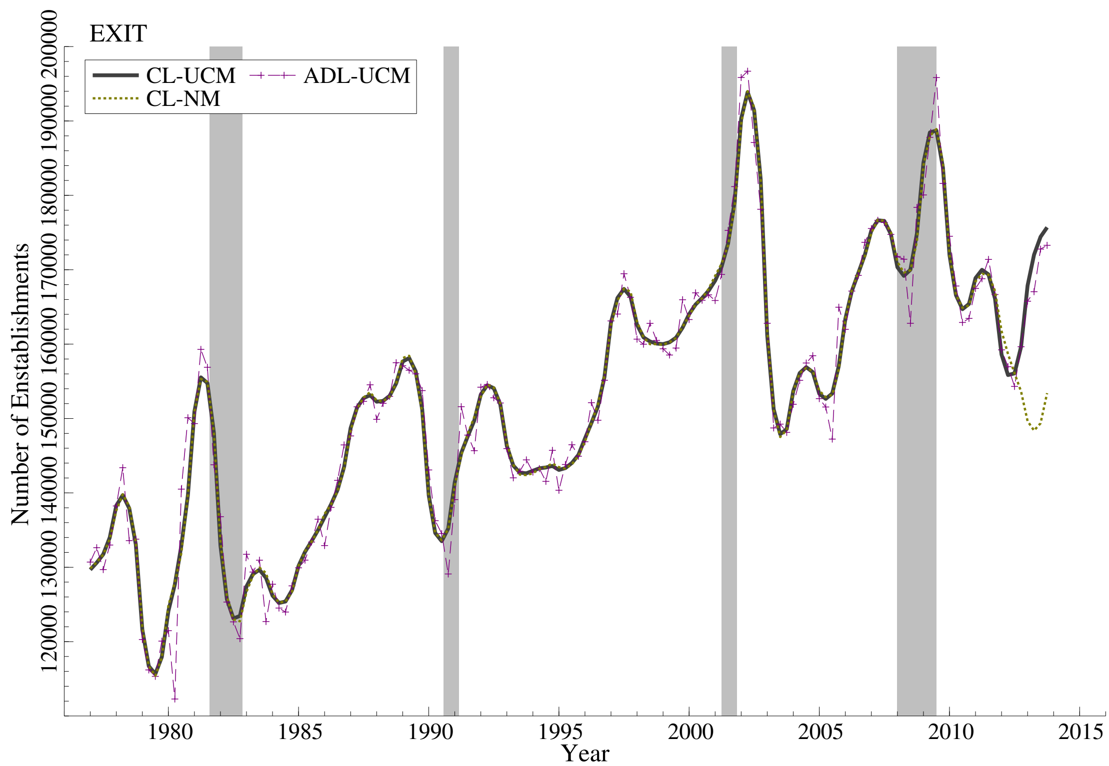

NOTE: This figure plots the cross-correlation function of the BLS data (OPENINGs and CLOSINGs) with RGDP; both the data are in growth rates. Red values indicates leads while blue value indicates lags. The two horizontal lines are the critical values for the significance of the function.

**Figure 2:** The disaggregated series

**(a)** ENTRY



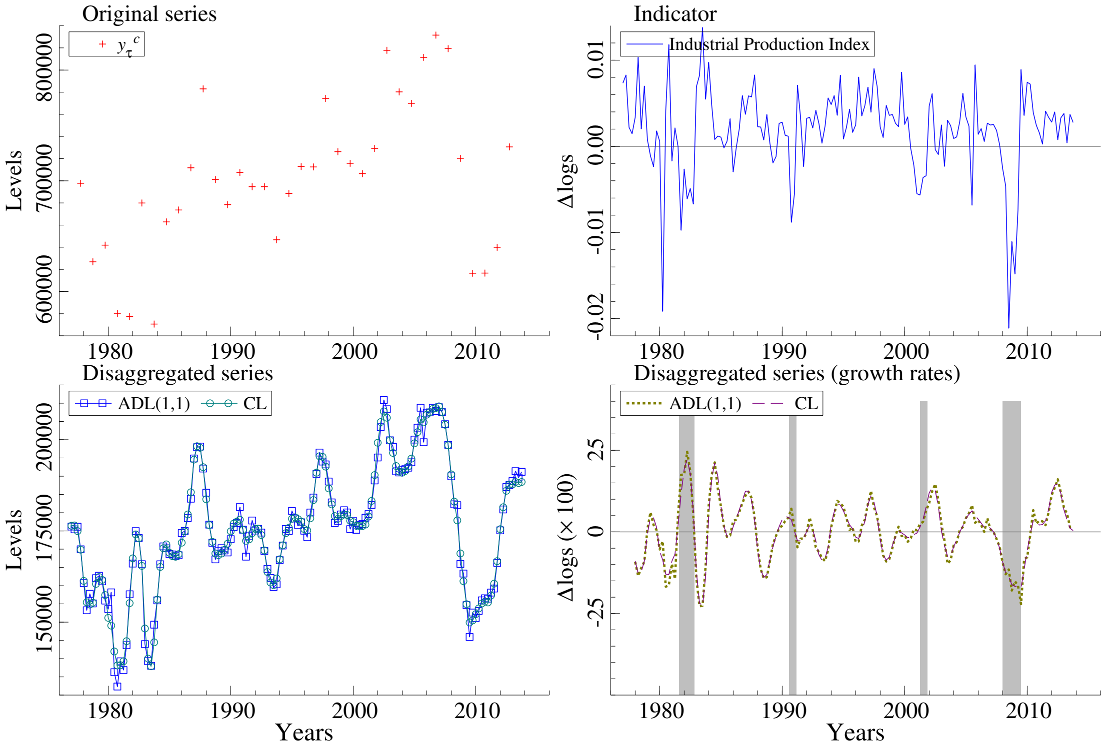

**(b)** EXIT



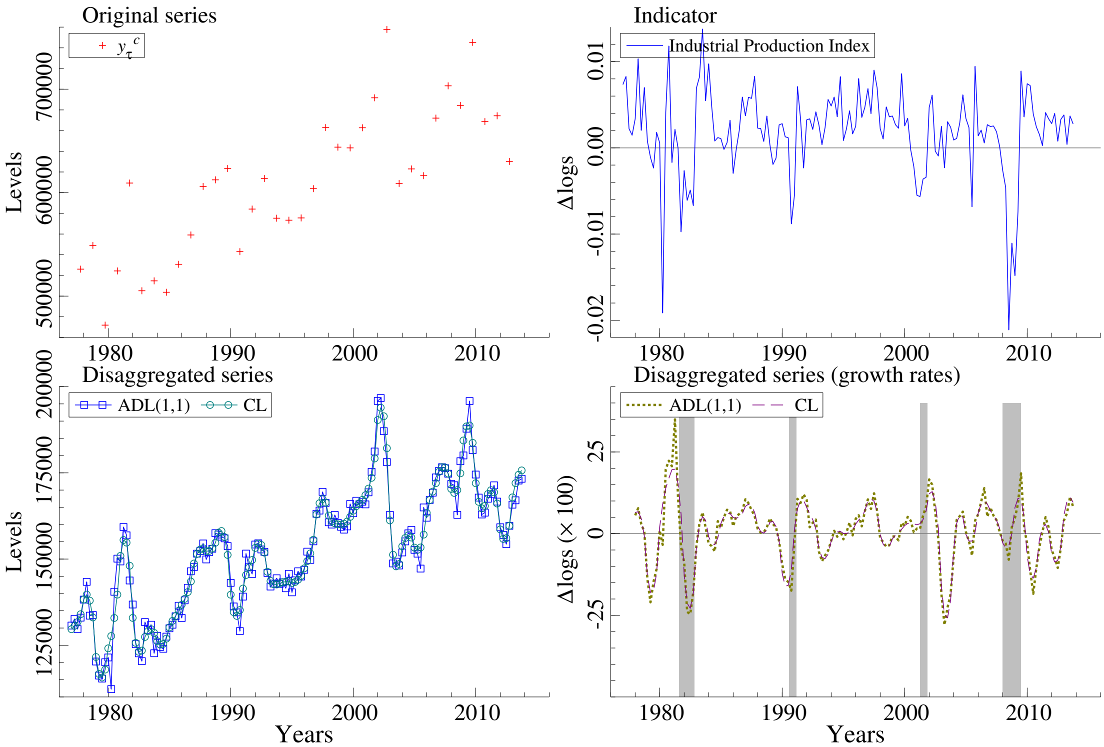

NOTE: This figure plots the values of the time series resulting from the three different temporal disaggregation methods exposed in Section 3. Colored bands correspond to the NBER recessions. Software used: OxMetrics

**Figure 3:** Comparison of different regression models of univariate UCM

**(a)** ENTRY

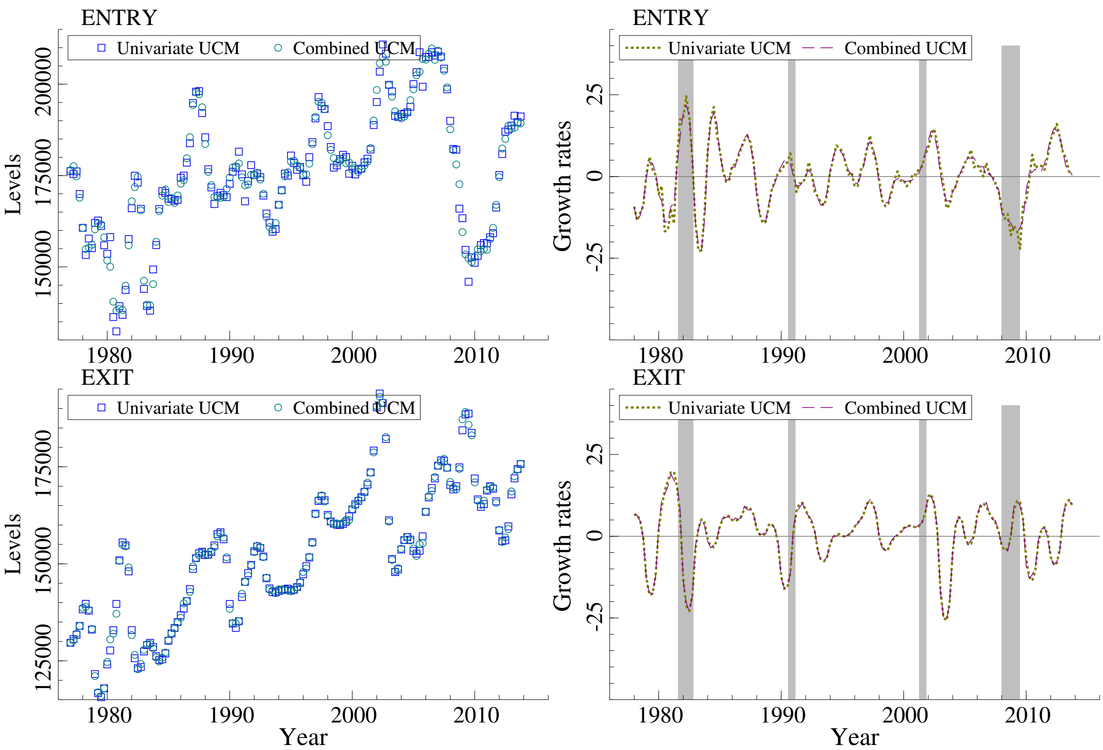



**(b)** EXIT



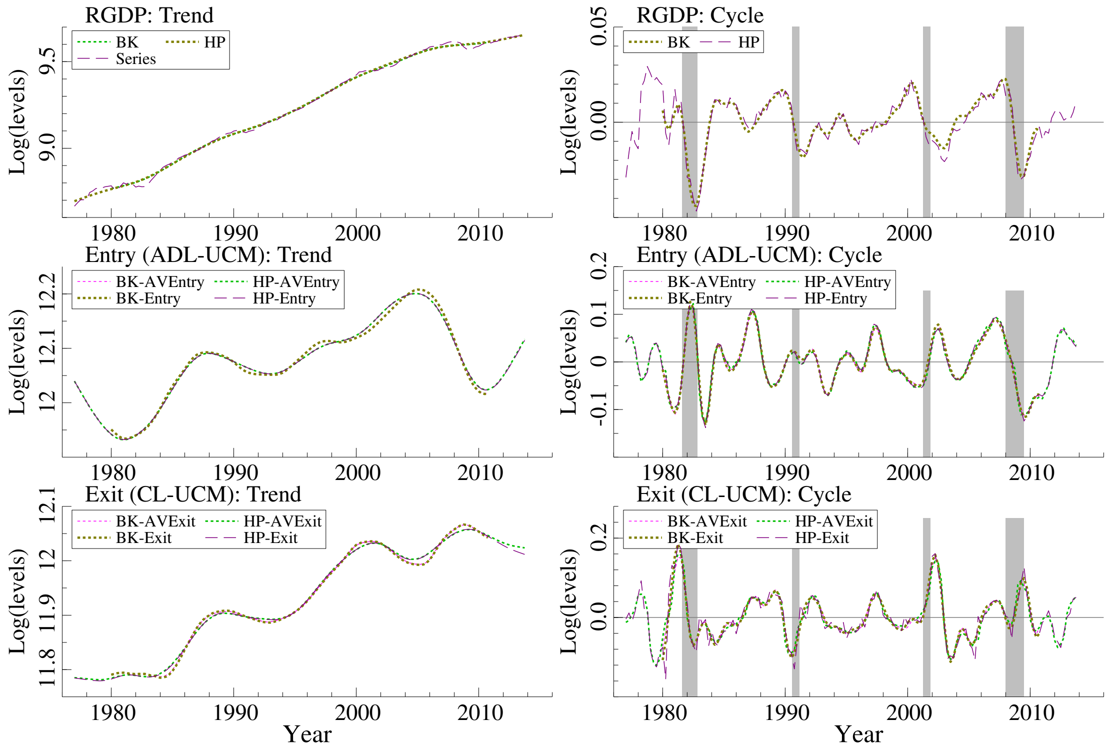

NOTE: This figure plots the ingredients forming the univariate UCM discussed in Section 3.2, and namely: (i) the original yearly-frequency series (upper left panel); (ii) the indicator variable (upper-right); (iii) the resulting disaggregated series according to ADL and CL models of UCM (lower-left); and (iv) the same disaggregated series, but in growth rates (lower-right). Colored bands correspond to the NBER recessions. Software used: OxMetrics

**Figure 4:** Comparison of univariate and combined UCM.




NOTE: This figure compares the series resulting from the application of univariate disaggregation discussed in Section 3.2 and combined UCM discussed in Section 3.2.3. The models are selected according to the results of LR-test in Table 2. Upper panels refers to ENTRY and lower panels to EXIT; on the other side, left panels plot series in levels and right panels plot the same series in growth rates. Colored bands correspond to the NBER recessions. Software used: OxMetrics.

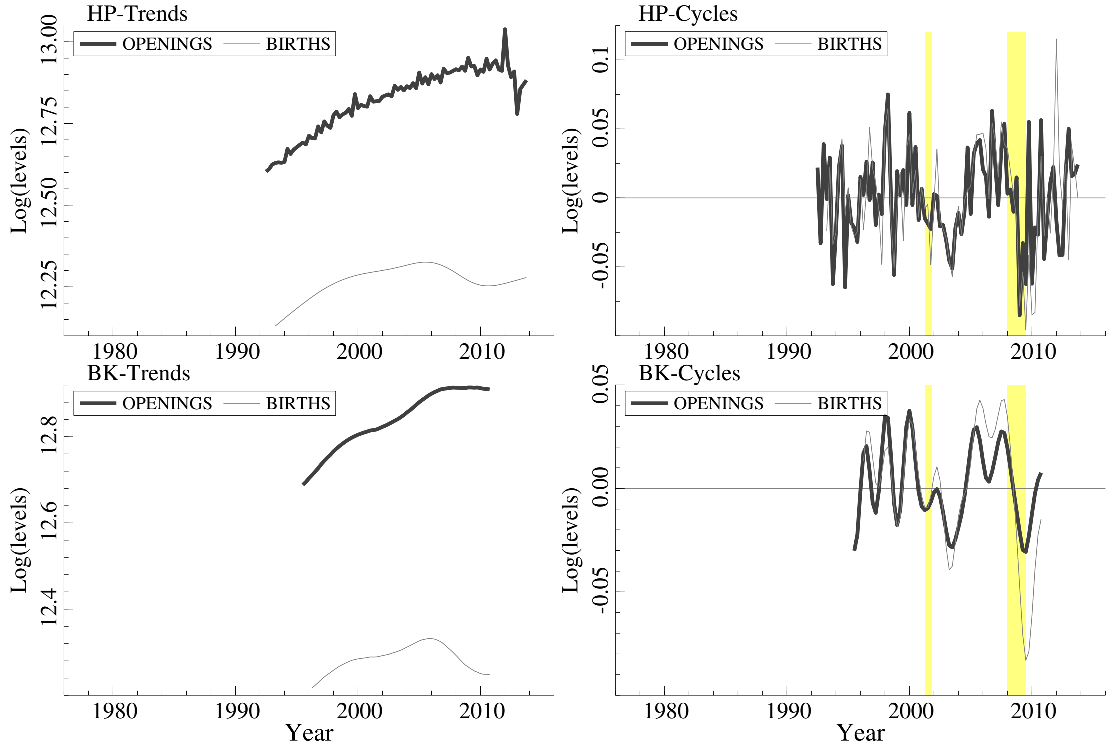
**Figure 5:** Business cycle analysis of disaggregated series



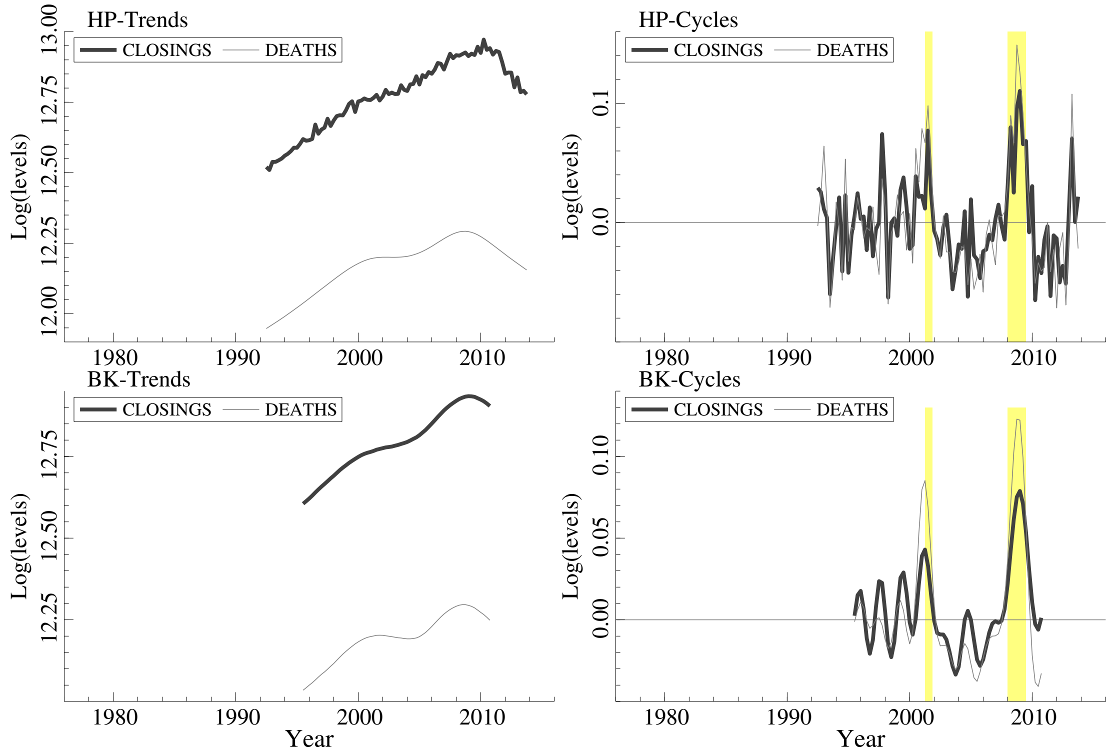

NOTE: This figure shows the business cycle analysis of the RGDP (upper panel) and the ENTRY and EXIT series (central and lower panel, respectively) resulting from the application of univariate disaggregation discussed in Section 3.2 and combined UCM discussed in Section 3.2.3. Namely, left (right) panels plot the trend (cycle) component extracted from the series in logarithms. The cyclical components extracted by the HP filter assume penalization parameter $\lambda = 1600$ and the ones obtained by BK filter assume parameters $\lambda_0 = 1.5$, $\lambda_1 = 8$ and $K = 12$ (that is, $3 \times s$). Colored bands correspond to the NBER recessions. Software used: OxMetrics.

**Figure 6:** Business cycle analysis of BLS quarterly series

**(a)** ENTRY



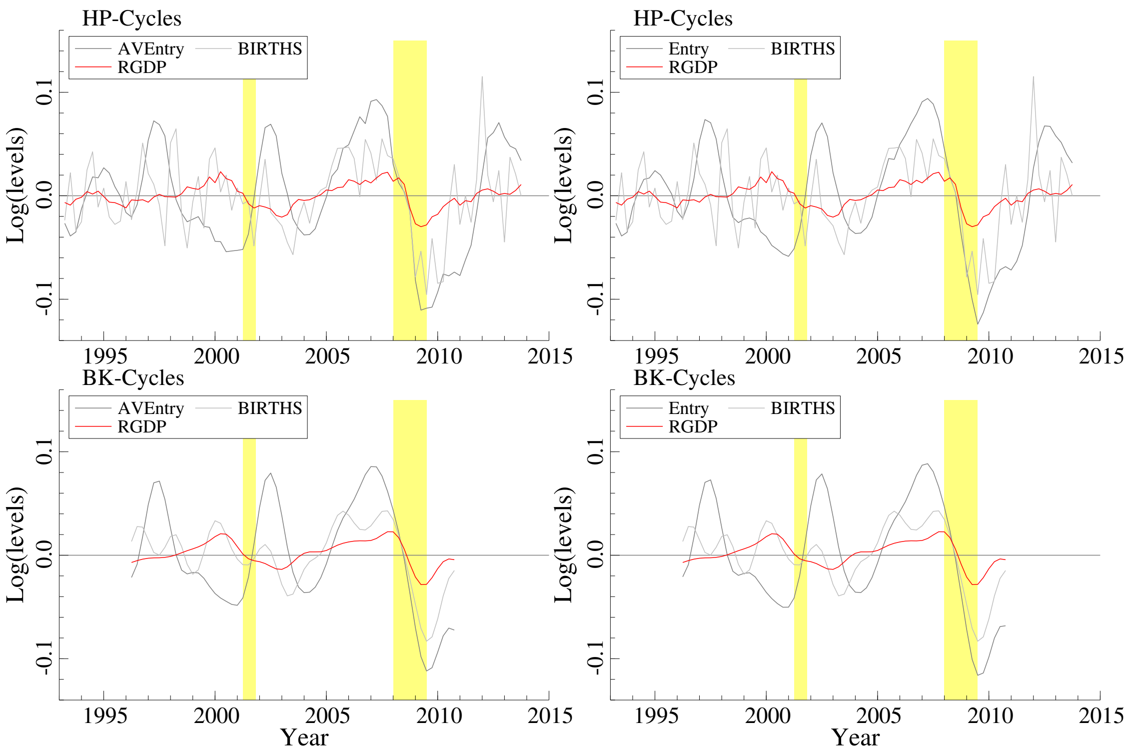

**(b)** EXIT



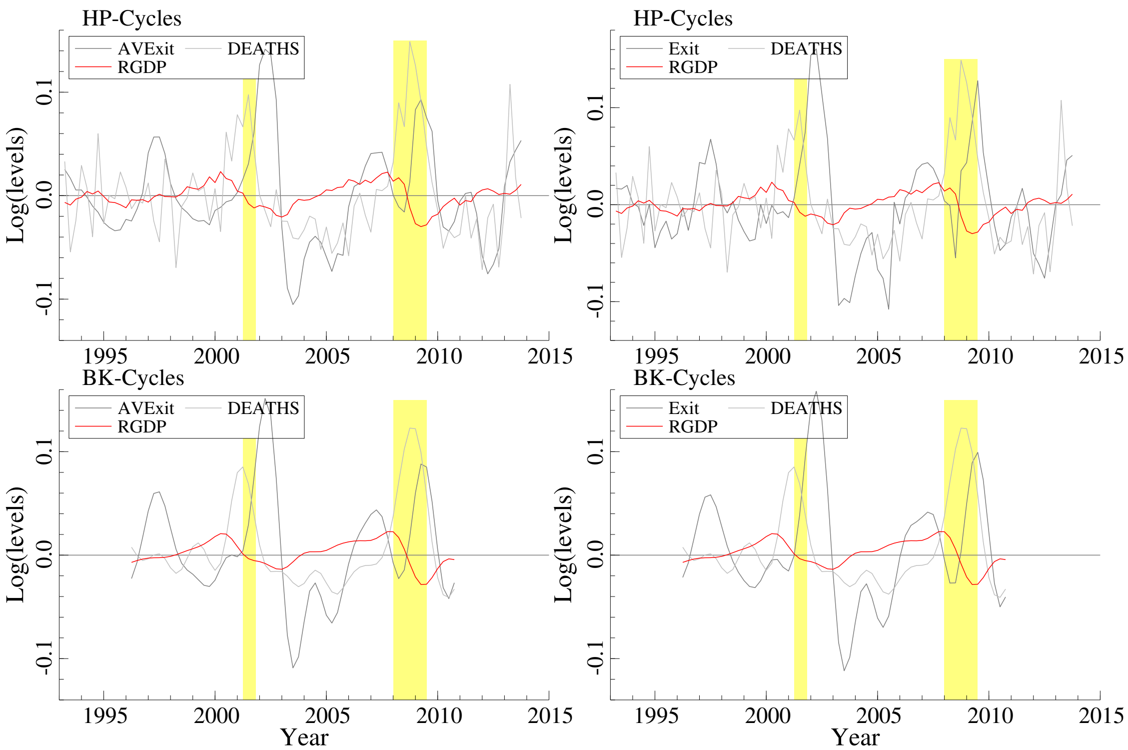

NOTE: This figure shows the business cycle analysis of the BLS existing quarterly series of ENTRY – in panel (a) – and EXIT – in panel (b). Left (right) panels plot the trend (cycle) component extracted from the series in logarithms. Upper (lower) panels refer to HP (BK) filter. The cyclical components extracted by the HP filter assume penalization parameter $\lambda = 1600$ and the ones obtained by BK filter assume parameters $\lambda_0 = 1.5$, $\lambda_1 = 8$ and $K = 12$ (that is, $3 \times s$). Colored bands correspond to the NBER recessions. Software used: OxMetrics.

**Figure 7:** The cycle component: comparison between different proxies.

**(a)** ENTRY

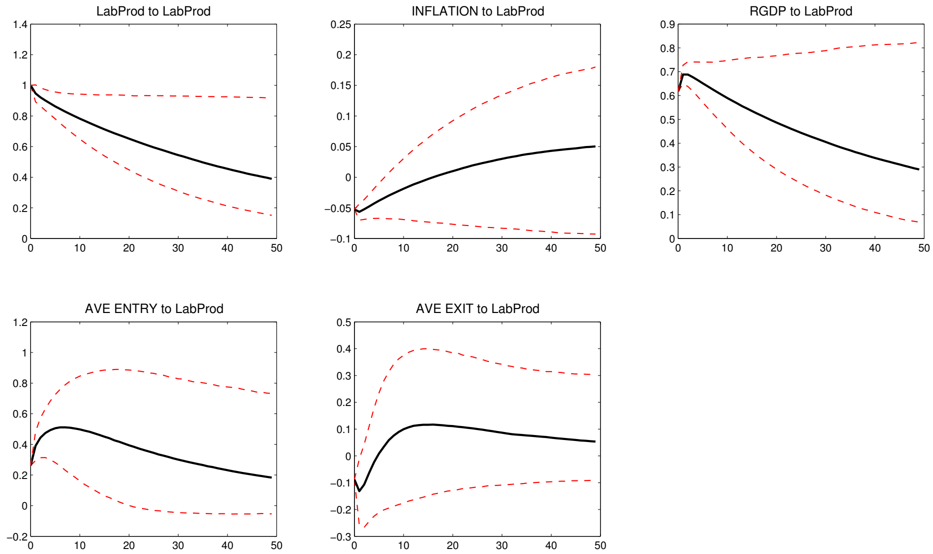



**(b)** EXIT



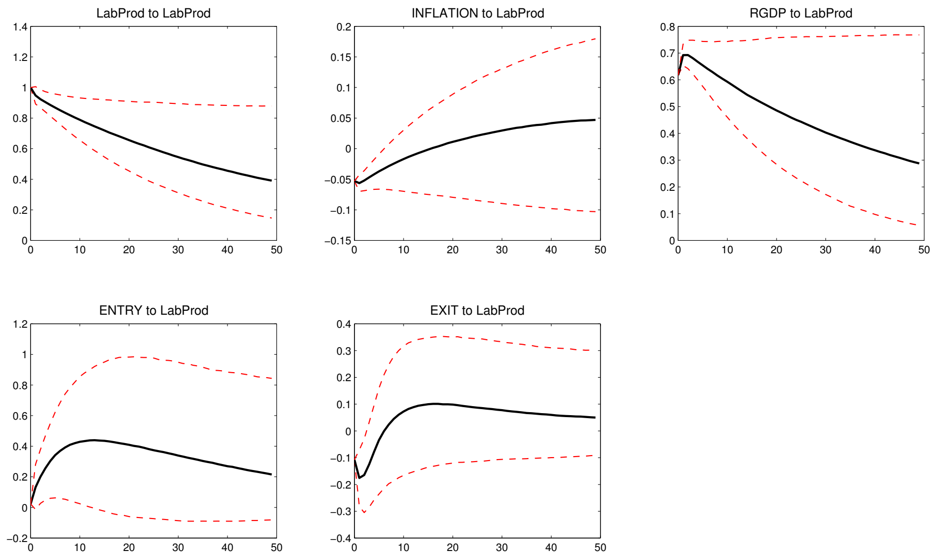

NOTE: This figure shows the cycle of BLS proxies of ENTRY – in panel (a) – and EXIT – in panel (b) for different filters. Namely, the upper (lower) panels refer to the HP (BK) filter. On the other side, left (right) panels compares these proxies with combination (univariate) measures of ENTRY/EXIT. The cyclical components extracted by the HP filter assume penalization parameter $\lambda = 1600$ and the ones obtained by BK filter assume parameters $\lambda_0 = 1.5$, $\lambda_1 = 8$ and $K = 12$ (that is, $3 \times s$). Colored bands correspond to the NBER recessions. Software used: OxMetrics.

**Figure 8:** IRFs from BVAR-M1 with Minnesota Priors.



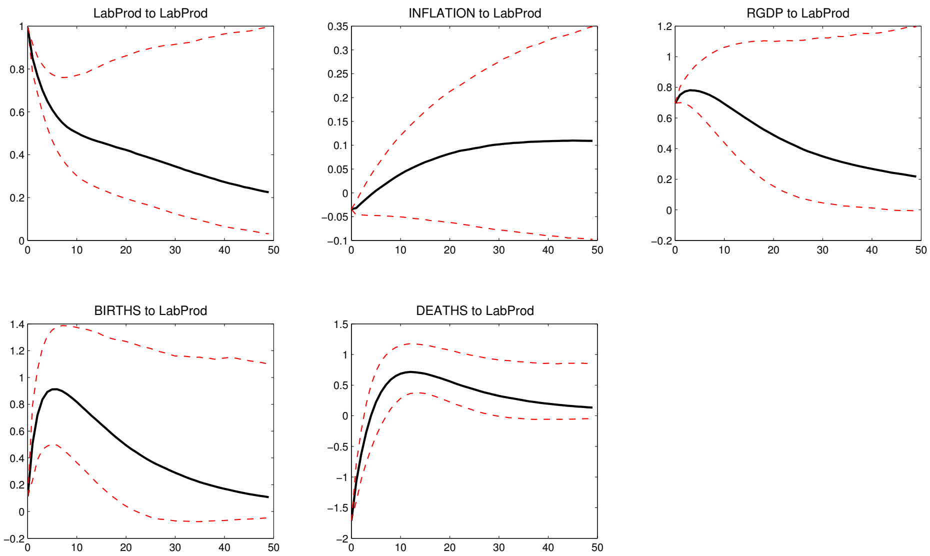

NOTE: This figure shows the responses of the BVAR-M1 system to a labor productivity shock of one standard deviation. Sample: 1983Q1–2013Q4.

**Figure 9:** IRFs from BVAR-M2 with Minnesota Priors.



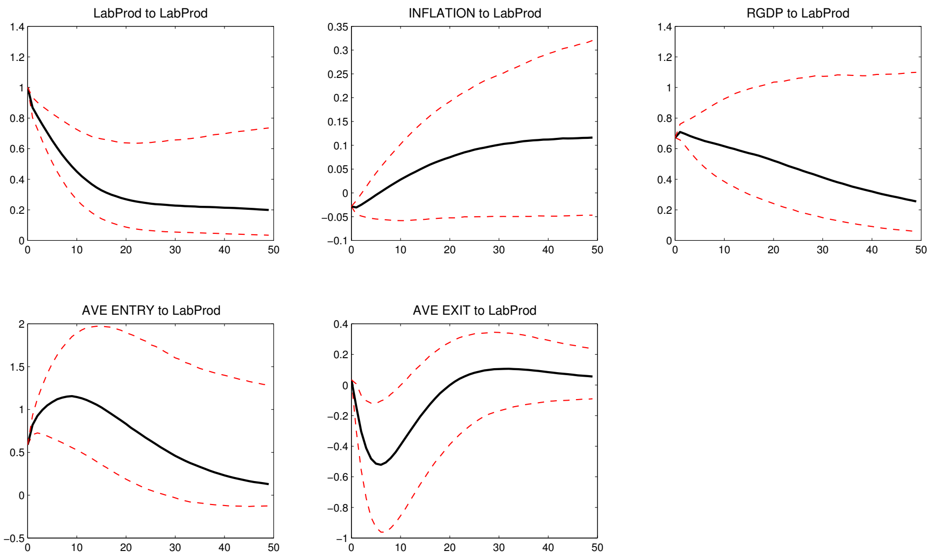

NOTE: This figure shows the responses of the BVAR-M2 system to a labor productivity shock of one standard deviation. Sample: 1983Q1–2013Q4.

**Figure 10:** IRFs from BVAR-M3 with Minnesota Priors.



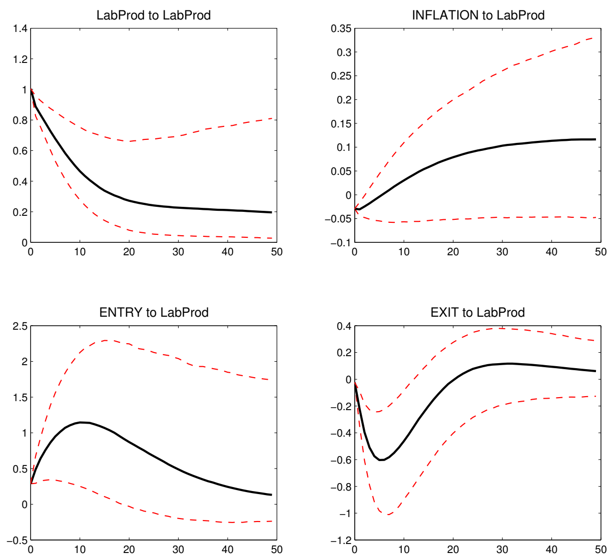

NOTE: This figure shows the responses of the BVAR-M3 system to a labor productivity shock of one standard deviation. Sample: 1993Q3–2013Q4.

**Figure 11:** IRFs from BVAR-M4 with Minnesota Priors.



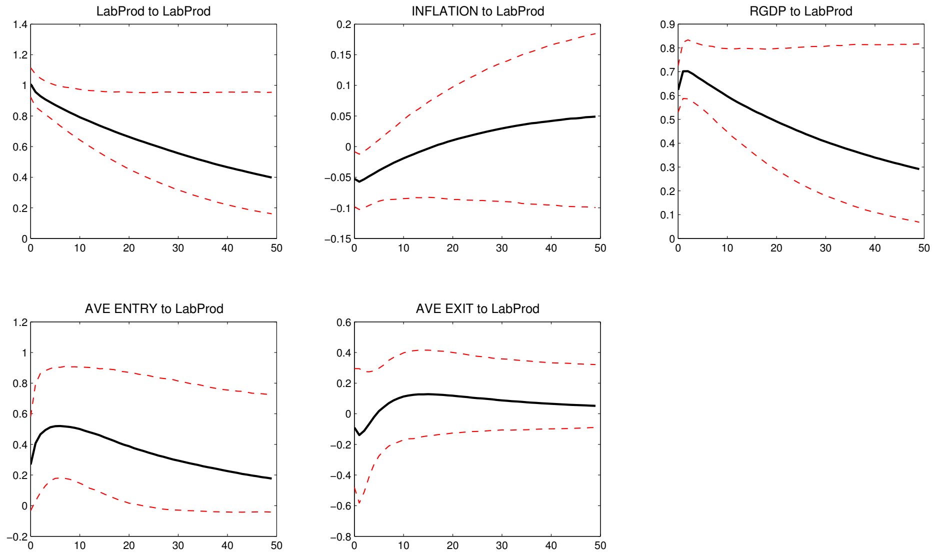

NOTE: This figure shows the responses of the BVAR-M4 system to a labor productivity shock of one standard deviation. Sample: 1993Q3–2013Q4.

**Figure 12:** IRFs from BVAR-M5 with Minnesota Priors.



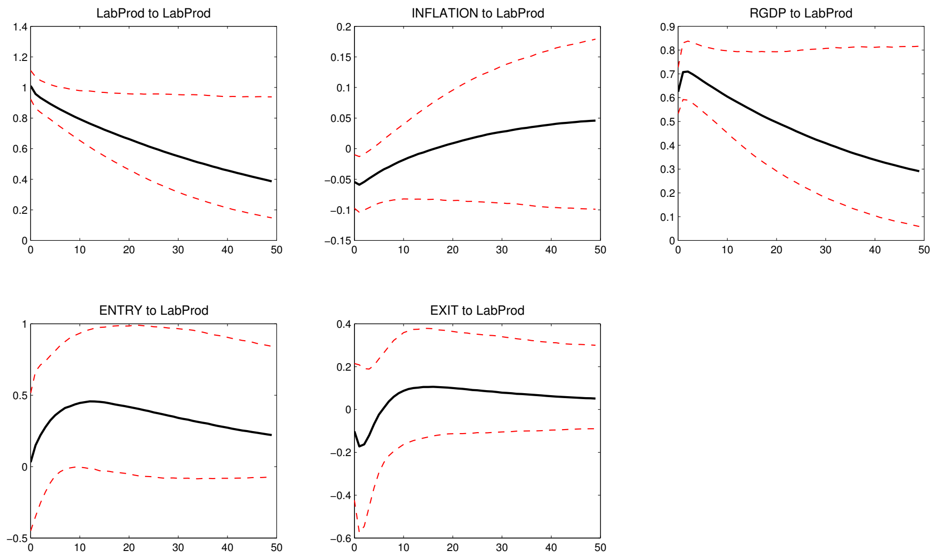

NOTE: This figure shows the responses of the BVAR-M5 system to a labor productivity shock of one standard deviation. Sample: 1993Q3–2013Q4.

**Figure 13:** IRFs from BVAR-M1 with Normal Diffuse Priors.



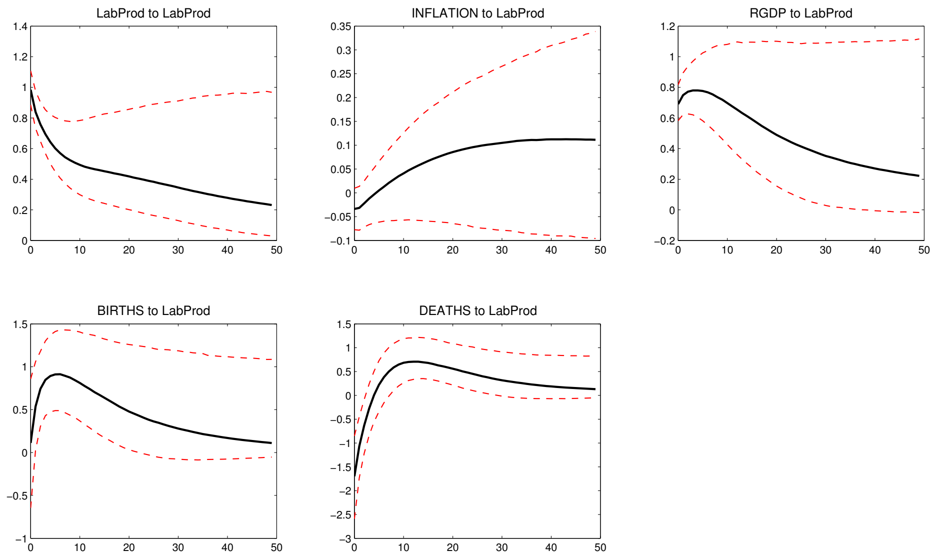

NOTE: This figure shows the responses of the BVAR-M1 system to a labor productivity shock of one standard deviationt. Sample: 1993Q3–2013Q4.

**Figure 14:** IRFs from BVAR-M2 with Normal Diffuse Priors.



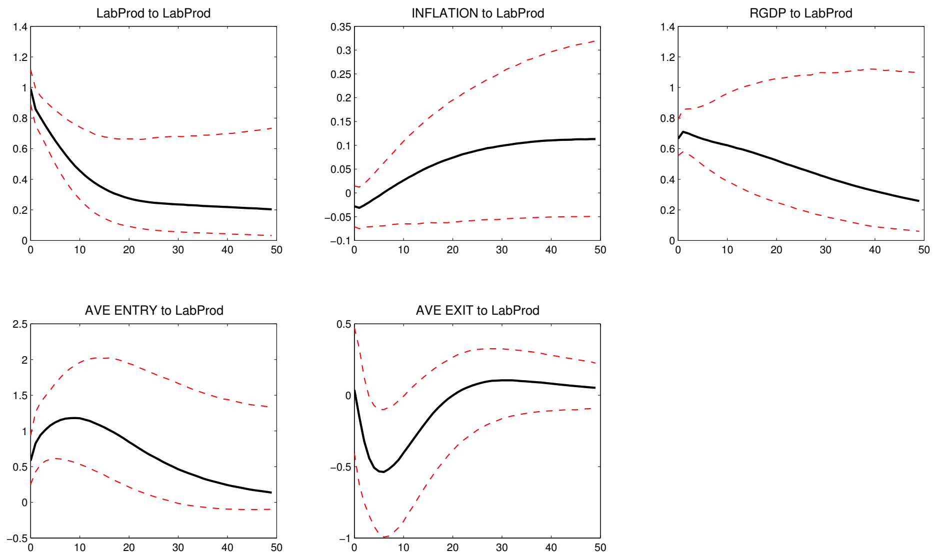

NOTE: This figure shows the responses of the BVAR-M2 system to a labor productivity shock of one standard deviation. Sample: 1983Q1–2013Q4

**Figure 15:** IRFs from BVAR-M3 with Normal Diffuse Priors.



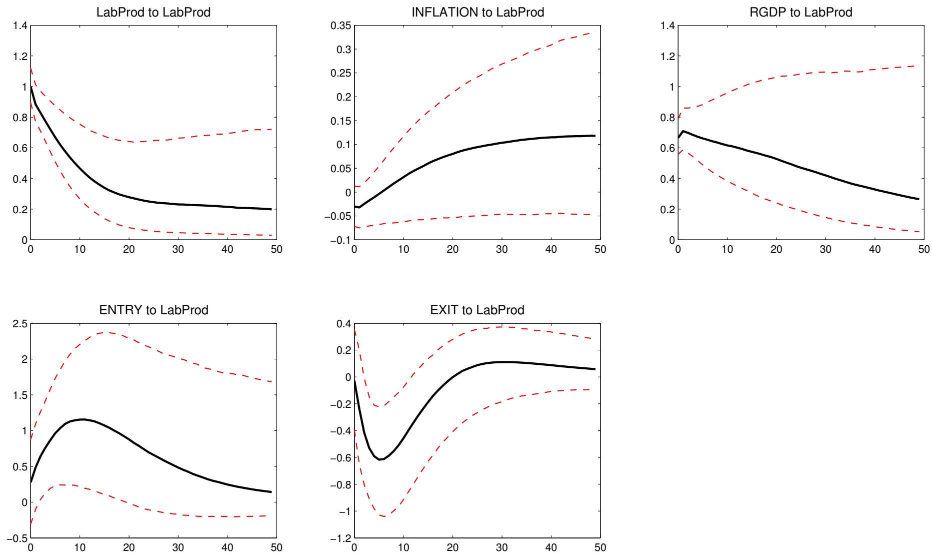

NOTE: This figure shows the responses of the BVAR-M3 system to a labor productivity shock of one standard deviation. Sample: 1993Q3–2013Q4

**Figure 16:** IRFs from BVAR-M4 with Normal Diffuse Priors.




NOTE: This figure shows the responses of the BVAR-M4 system to a labor productivity shock of one standard deviation. Sample: 1993Q3–2013Q4.

**Figure 17:** IRFs from BVAR-M5 with Normal Diffuse Priors.




NOTE: This figure shows the responses of the BVAR-M5 system to a labor productivity shock of one standard deviation. Sample: 1993Q3:2013Q4

**Table 1:** Correlations between disaggregated series and BLS data

| Method | Corr(ENTRY, OPENINGS) | Corr(EXIT, CLOSINGS) | SAMPLE |
|---|---|---|---|
| CL-NM | 0.1625 | 0.5279 | |
| CL-UCM | 0.2669 | 0.7171 | 1992:Q3–2013:Q4 |
| ADL-UCM | 0.2703 | 0.7225 | |
| | Corr(ENTRY, BIRTHS) | Corr(EXIT, DEATHS) | SAMPLE |
| CL-NM | 0.1792 | 0.7137 | |
| CL-UCM | 0.1913 | 0.7322 | 1993:Q2–2013:Q4 |
| ADL-UCM | 0.2260 | 0.7152 | |

NOTE: This table reports the correlation between the series resulting from disaggregation method accrding to (1) and the method (8) with the specifications (3) and (4) and the existing quarterly data downloaded from BLS.

**Table 2:** Estimated model parameters of UCM for different specifications

| REGRESSION MODEL | PARAMETER | | ENTRY | | | EXIT | | |
|---|---|---|---|---|---|---|---|---|
| ADL(1,1) | Log-Likelihood | | -437.6925 | | | -431.91524 | | |
| | $\phi$ | | 0.8045 | | | 0.6200 | | |
| | $\sigma^2$ | | $8.9847e^7$ | | | $-1.0435e^8$ | | |
| | | | Value | StDev | t-stat | Value | StDev | t-stat |
| | Regression Effects | $x_1$ | 30,716.87 | 1,599.51 | 19.15 | 49,024.53 | 1,780.91 | 27.53 |
| | | $x_2$ | 41.89 | 18.15 | 2.31 | 119.12 | 20.20 | 5.90 |
| | | $x_3$ | -409,716.78 | 533,143.27 | -0.77 | 689,839.94 | 552,907.00 | 1.25 |
| | | $x_4$ | 769,250.14 | 515,659.34 | 1.49 | | | |
| Chow-Lin | Log-Likelihood | | -439.8661 | | | -431.9442 | | |
| | $\phi$ | | 0.8015 | | | 0.4990 | | |
| | $\sigma^2$ | | $-1.0228e^8$ | | | $-1.4756e^8$ | | |
| | | | Value | StDev | t-stat | Value | StDev | t-stat |
| | Regression Effects | $x_1$ | 32,623.80 | 1,556.93 | 20.95 | 64,365.34 | 2,077.61 | 30.98 |
| | | $x_2$ | 33.73 | 18.90 | 1.78 | 163.93 | 24.08 | 6.81 |
| | | $x_3$ | -143,156.25 | 552,615.74 | -0.26 | 36,358.01 | 479,154.96 | 0.08 |
| LR-statistic | | | **4.3471** | | | 1.5835 | | |

NOTE: This table reports the estimates of parameters of equations (3) and (4). Regressors $x_1, \ldots, x_4$ correspond to the deterministic component included in the model, see Remark 2; all the diffuse effects $x_t$ are reported, jointly with their standard deviation and t-statistics. The log-Likelihood corresponds to diffuse profile (14) deriving from estimated parameters in (12). The one-degree-of-freedom Likelihood-ratio test corresponds to the null hypothesis that $\phi = 0$ versus the alternative hypothesis that $\phi \neq 0$ in condition (5); rejection of null hypothesis leads to ADL(1,1) model; the bold is used to indicate that statistic is rejected at 10% of significance.

**Table 3:** Stylized Business Cycle Facts

| Data Types | Series | Lead | Lag | Sign |
|---|---|---|---|---|
| | Entry-BK | | 3 | + |
| | Entry–HP | | 3 | + |
| | Entry–LDiff | 8 | | + |
| Disaggregated (Univariate UCM) | | | | |
| | Exit–BK | 6 | | - |
| | Exit–HP | | 8 | + |
| | Exit–LDiff | | 1 | - |
| | AVEntry-BK | | 3 | + |
| | AVEntry–HP | | 3 | + |
| | AVEntry–LDiff | | 1 | + |
| Disaggregated (Combined UCM) | | | | |
| | AVExit–BK | 6 | | - |
| | AVExit–HP | 7 | | - |
| | AVExit–LDiff | 7 | | - |
| | Openings-BK | | 0 | + |
| | Openings–HP | | 0 | + |
| | Openings–LDiff | | 0 | + |
| BLS | | | | |
| | Closings–BK | | 5 | + |
| | Closings–HP | | 5 | + |
| | Closings–LDiff | | 4 | + |
| | Births-BK | | 0 | + |
| | Births–HP | | 0 | + |
| | Births–LDiff | | 1 | + |
| BLS | | | | |
| | Deaths–BK | 4 | | - |
| | Deaths–HP | 5 | | + |
| | Deaths–LDiff | | 2 | + |

NOTE: This table reports the cross-correlation analysis of the cyclical component of different measure of ENTRY and EXIT respect to the RGDP. "Lead" ("Lag") indicates the point where the cross-correlation function – in absolute value – of ENTRY (EXIT) and RGDP reach its maximum; a zero value of lag/lead indicate that the series is a coincident indicator of the business cycle. The sign "+" ("-") means that the maximum cross-correlation in the indicated lead/lag is positive (negative). The cyclical components extracted by the HP filter assume penalization parameter $\lambda = 1600$ and the ones obtained by BK filter assume parameters $\lambda_0 = 1.5$, $\lambda_1 = 8$ and $K = 12$ (that is, $3 \times s$).

**Table 4:** BVAR Models description

| Model | Sample | Priors | $Y_t$ |
|---|---|---|---|
| M1 | 1983:Q1–2013:Q4 | Minnesota | [LLabProd LCPI LRGDP LAVEntry LAVExit ]' |
| M2 | | | [LLabProd LCPI LRGDP LEntry LExit ]' |
| M3 | 1993:Q2–2013:Q4 | Minessota | [LLabProd INF DLRGDP BIRTHS DEATHS ]' |
| M4 | | | [LLabProd LCPI LRGDP LAVEntry LAVExit ]' |
| M5 | | | [LLabProd LCPI LRGDP LEntry LExit ]' |
| M1b | 1983:Q1–2013:Q4 | Normal Diffuse | [LLabProd LCPI LRGDP LAVEntry LAVExit ]' |
| M2b | | | [LLabProd LCPI LRGDP LEntry LExit ]' |
| M3b | 1993:Q2–2013:Q4 | Normal Diffuse | [LLabProd LCPI LRGDP LBIRTHS LDEATHS ]' |
| M4b | | | [LLabProd LCPI LRGDP LAVEntry LAVExit ]' |
| M5b | | | [LLabProd LCPI LRGDP LEntry LExit ]' |

NOTE: This table reports a list of definitions of vector $\mathbf{y_t}$ to be modeled by the BVAR system (16) for structural analysis. Namely, M1–M5 are BVAR estimated using Minnesota Priors. The labels M1b-M5b indicate that the BVAR models considering the Normal Diffuse Priors.