SPOUDAI Journal of Economics and Business, Vol.67 (2017), Issue 1, pp. 22-37

University of Piraeus

SPOUDAI
Journal of Economics and Business
Σπουδαί
http://spoudai.unipi.gr

A Generalized Dynamic Factor Model for the U.S. Port Sector

Jason Angelopoulos[a], Costas I. Chlomoudis[b]

[a] *University of Piraeus, Department of Maritime Studies, Greece.*
Email: jason.angelopoulos@gmail.com
[b] *University of Piraeus, Department of Maritime Studies, Greece.*
Email: chlom@unipi.gr

Abstract

Although rarely available ports produce a polymorphic set of timely available monthly import, export, transport and labor utilization series, providing frequent snapshots of freight volumes either as being transferred between modes, or trans-shipped to secondary destinations. Utilizing monthly inflows and outflows of several cargo types as well as cruise passenger volumes from U.S ports, we: a) demonstrate the potential and the added value of information carried by common factors shaped by ports with respect to outlining the underlying forces of a national economy and b) provide competitive forecasts of disaggregate trade series from single ports (such as, e.g. outgoing or incoming TEUs) by exploiting factor dynamics, We test this concept in the context of Forni et al. (2005) one-sided generalized dynamic factor model, exploring the links between ports and the driving factors of the U.S. economy, as these are captured through its common and idiosyncratic components. Our model, employing 192 series from 31 major port complexes -covering 84.4% of TEUs and 60.1% of the dry bulk volume between 2005 and 2012-, displays a promising forecasting performance for individual ports and aggregate economic indicators versus benchmark models at 4-7 months ahead and explains a high fraction of the US GDP and Industrial Production indices variance.

Keywords: dynamic factor models, U.S ports, trade, forecasting
JEL Classification: C38, F47, L99

1. Introduction

The research field of port services in general and port economics in particular, is constantly growing in the last 15 years; the main reasons behind this growth include efficiency requirements, management of operational complexity, and the evolution of governance structures and ownership schemes. Seaports constitute an integral part of transportation and their investments are affecting the allocation of transfer capacity (Chlomoudis and Pallis 1996, 1997). For any given domestic or international trade route, as long as waterborne transportation is more efficient or effective than other modes, ports will be a less replaceable part of the supply chain, since they are more likely to be competed by other ports than by other modes. Although no formal reporting process exists for U.S ports (National Cooperative Freight Research Program. 2014. p.126), they can produce a plethora of timely available monthly import, export, transport, port labor utilization series, as well as passenger volumes, providing frequent snapshots of freight volumes either as being transferred between modes, or trans-shipped to secondary destinations. Port passenger and trade flows, depending on geography, represent an interface where national and international production inputs / outputs and transportation, meet. Our intention in this exercise is to shift the point of view from the service industry aspect of ports, bilateral or multilateral trade issues, or economic interpretation of aggregate trade or transport indexes, in order to outline their properties as regional and global facilitators between demand and supply and demonstrate their capacity for reflecting national economic activity and forecasting. Following the emergence of the U.S transportation sector (Lahiri, 2010) as one of bearing coincident and leading added value related to the national U.S. economy and given the importance of the port sector for national transportation, production, and international trade flows, it is worth exploring possible relationships and co-movements of previously unexploited monthly port inputs and outputs both with each other, as well as with aggregate economic indexes.

The current study attempts to contribute to the convergence of contemporary port literature and factor analysis of economic fluctuations in the context of a national economy, as well as macroeconomic and operational forecasting: By utilizing a large national port dataset from the U.S., we evaluate forecasts of macroeconomic variables, using ports as instruments, and forecasts of individual ports, using the rest of the ports as instruments. This is an exercise well suited for a factor model, which main input are large co-moving datasets, with the outputs being a limited number of factors.

The reduction of dimensionality achieved by a small number of common factors facilitates the assessment of co-movements of large cross-sections in the economy and exploits its common components for forecasting, has been demonstrated by various applications of macroeconomic factor models on a national or even multi-national level the last 15 years. Therefore, the idea of extending this concept to a sector-specific polymorphic set of disaggregate data is both technically feasible, as well as plausible. To our knowledge this is the first application of a large-scale generalized dynamic factor model assessing information originating from the U.S port sector, and the second, after Reijer (2007), that deals with exclusively disaggregate data from any sector on a national level. Utilizing monthly inflows and outflows of several cargo types, port employment and cruise passenger volumes from the majority of U.S ports, we cover 84.4% of TEUs and 60.1% of the dry bulk volume from 2005 to 2012, and for the first time a) we demonstrate the close co-movement of disaggregate port series with the drivers of the U.S economy and b) we provide evidence for their forecasting capabilities with respect to individual ports, specific cargo types, cruise passengers, port workforce requirements and aggregate economic indicators.

In principle, disaggregate data can be characterized as driven mainly by idiosyncratic disturbances and present a more volatile nature. But the possibility that aggregation may cancel out opposite trends or dampen converse co-movements cannot be ruled out. Using the same argument, the possibility of useful information being present only in partially or fully disaggregate data cannot be ruled out either. By utilizing the same econometric methods and theoretical background we are able to explore the answers to two questions: a) what is the potential and the added value of information carried by common factors shaped by ports with respect to outlining the underlying forces of a national economy.

b) whether we can provide reliable forecasts of disaggregate trade, passenger and workforce utilization series from single ports (such as, e.g. outgoing or incoming TEUs), by exploiting factor dynamics shaped from all other ports, and

In the next section we provide background information about relevant literature and we explain the methodology utilized. In section 3 we discuss factor model parameter selection, documenting our parameter choices; in section 4 we discuss port panel dynamics and compare forecasts of our model vs. benchmarks and in section 5 we summarize our main findings.

2. Background

Factor models, in general, are tools to facilitate parsimonious representation of large cross sections of time series by a small number of exogenous common shocks shared by panel series, which can be used also for forecasting. They are utilized as non-structural models and macroeconomic forecasting tools by central banks, governments and academia. Major developments of factor models in economic applications in the last 15 years can be found in Stock and Watson (2002) and Forni et al. (2000), refining the concepts of so-called static factor and two-sided dynamic factor models respectively and detailing their application in economic series. Several enhancements and alternatives were proposed, most notably Forni et al., (2005), Kapetanios and Marcellino (2006) and Doz et al., (2011) utilizing one-sided filtering, state-space methods and the Kalman filter respectively. A comprehensive review can be found in Stock and Watson (2011). Applications on a national economy level followed shortly after Stock and Watson (2002) and Forni et al. (2000); Notable examples of dynamic or static implementations include Schumacher and Dreger (2002), Nieuwenhuyze (2006), Banerjee and Marcellino (2006), Carriero and Marcellino (2007), Cheung and Demers (2007), Reijer (2007), Ajevskis and Dāvidsons (2008), Nguiffo-Boyom (2008), Barhoumi et al., (2010) for Germany, Belgium, U.S, U.K, Canada, Netherlands, Latvia, Luxemburg and France respectively. Factor models have been also utilized in a multi-national scale context in Altissimo et al., (2001) and Forni et al., (2001) for the Euro Area, Al-Hassan (2009) for the Gulf Cooperation Council and Guichard and Rusticelli (2011) for international trade growth. In most of the applications cited above, as well as in Barhoumi et al., 2010 and Schumacher (2007), dynamic and static factor models outperform naïve and autoregressive models in GDP forecasting; furthermore, depending on dataset attributes and forecast horizons, they display similar forecasting performance. Assessing the methodological differences between static and dynamic approaches, as well as comparing their forecasting accuracy is outside the scope of the current work. For this exercise we will utilize the Forni et al., (2005) one-sided generalized dynamic factor model, GDFM hereafter, mostly for its useful capability for assessing leads and lags between variables, enabling us to set macroeconomic variables as reference series for the U.S port sector.

The intuition that disaggregate trade data may be applicable for forecasting economic variables is not new. Perevalov and Meier (2010), focusing on U.S GDP forecasting with factor models, concluded that 'the largest improvements in terms of forecasting accuracy are found for relatively more volatile series, with the greatest gains coming from improvements of the forecasts for investment and trade'. In Altissimo et al., (2001) it is stated that shocks originating from a local or sectoral source generate dynamics that should be monitored by local or sectoral policy makers. In relation to the scope of this exercise, with the exception of Tsamourgelis et al., (2013), the link between trade and the economy has also been explored in the past, but not through a disaggregate polymorphic port set. In the context of bilateral trade we note Aruoba et al., (2010), which concludes that extent of synchronization between countries in the '00s has not changed since the 1970-1980 period despite the increase in global trade and financial linkages, as well as the robust correlation of business cycles and bilateral trade between two countries, assessed in Baxter and Kouparitsas (2004). We also find a useful stream of work about the U.S transportation services index (TSI) summarized in Lahiri (2010), where the potential predicative quality of freight transport is elaborated, and its closely connection with inventories is highlighted.

In GDFM, variable dynamics are virtually split into two mutually orthogonal components: a) the common component, a linear combination of all factors shared by each of the panel series with a varying degree of commonality and b) the idiosyncratic component, comprised of series-specific factors, disturbances and measurement errors. The observations in the generalized principal components are weighted based on their signal to noise ratio. The first step involves the calculation of the spectral density matrixes (frequency domain) and the autocovariances (time domain) for the common and idiosyncratic components. The second step includes the computation of the linear combinations that maximizes the contemporaneous covariance induced by the common factors. More specifically, the GDFM is estimated as follows:

X_{Nt} is defined as a panel of N series with $t = 1,2,...T$ observations. The process X_t is the sum of two unobservable components, the common factors χ_t (alternatively referred to as *common* or *primitive* shocks) and the idiosyncratic factors ξ_t, alternatively $X_t = \chi(F_t) + \xi_t$, where F_t is the lag operator for $q >> N$ common factors. $\chi(F_t)$ is a two-sided filter of X_t, implying a deterioration of forecasts as $t \to T$.

To circumvent this shortcoming, Forni et al., (2005) utilized the frequency domain by estimating the spectral density matrix, applied dynamic principal components analysis, and then reverted back to the time domain by inverse Fourier transform. Then, using the spectral density matrices, the covariance matrices for the common and idiosyncratic components are obtained for all leads and lags, smoothed over M frequencies, using generalized principal components. Finally, the estimated common components are projected orthogonally by the static factors. They represent the r contemporaneous linear combinations of x_t with the lowest idiosyncratic to common variance ratio and they reflect the degree of heterogeneity with respect to the reaction (impulse response) of each common factor. Utilizing the GDFM we can assess the extent of the variance explained by the variance of its common component:

X_{Nt} is explained by the variance Γ^{χ}_{Nk} of its common component χ_t and Γ^{ξ}_{Nk} the variance of its idiosyncratic component ξ_t. The total panel variance Γ^{T}_{Nk} is defined as:

$$\Gamma^{T}_{Nk} = E[X_{\mathrm{N}t}(X_{\mathrm{N}t-k})^{\mathrm{T}}] \tag{1}$$

where k are the number of lags, and $(.)^{\mathrm{T}}$ denotes transposition.

We compute autocovariance matrices (sample covariance) of order k ($-k, ..., 0, ..., k$), defined by:

$$\Gamma^{T}_{Nk} = (T-k)^{-1} \sum_{t=k+1}^{T} X_{Nt} (X_{Nt})^{\mathrm{T}} \ . \tag{2}$$

The spectral density matrix, over w_k Bartlett-lag windows $w_k = \dfrac{|k|}{M+1}$, is estimated by the Fourier transform:

$$\Sigma^{\mathrm{T}}_{\mathrm{N}}(\theta_s) = \sum_{k=-M}^{M} w_k \Gamma^{T}_{Nk} e^{-i\theta_s k} \tag{3}$$

where $\theta_s = \dfrac{2s\pi}{2\mathrm{M}+1}$, $s = -M, M+1, ...M$, M is an integer, $M = M(T)$.

The next step involves the decomposition of $\Sigma^{\mathrm{T}}_{\mathrm{N}}(\theta_s)$ into $\Sigma^{\chi\mathrm{T}}_{\mathrm{N}}(\theta_s)$ and $\Sigma^{\xi\mathrm{T}}_{\mathrm{N}}(\theta_s)$ by applying dynamic principle component analysis (Brillinger, 1981, ch. 9), essentially by computing the $\Sigma^{\chi T}_{N}$ matrices, but using only the first q dynamic factors:

$$\Sigma^{\chi T}_{N} = \lambda^{T}_{N1}(\theta) p^{T}_{N1} + ... + \lambda^{T}_{Nq}(\theta)(p^{T}_{Nq})^{*} p^{T}_{Nq} \tag{4}$$

where $\lambda_{Nq}^{T}(\theta)$ denotes the largest eigenvalue of $\Sigma_{N}^{\chi T}$, p_{Nq}^{T} is the largest eigenvector and $(.)^{*}$ denotes conjugate transpose.

We will later assess the steps of choosing the 'right' number of q and M. We then revert into the time domain utilizing the inverse Fourier transform:

$$\Gamma_{Nh}^{\chi T} = (2M+1)^{-1} \sum_{h=-M}^{M} \Sigma_{N}^{\chi T}(\theta_s) e^{i\theta_s k} . \tag{5}$$

For each static factor r, the variance of the idiosyncratic factors is the residual variance, if we deduct the variance of χ_t at $M = 0 : \Gamma_{j0}^{\xi T} = \Gamma_{j0}^{T} - \Gamma_{j0}^{\chi T}$, where $j \in [1,2,...,r]$. For reference, in the empirical application of Forni et al. (2005), a range of 6 to 15 static factors is utilized.

Finally, the step of calculating the generalized principle components K_{N}^{Th} is performed as follows: Denoting Z_{Nj}^{T} as the generalized eigenvectors matrix for $\Gamma_{j0}^{\chi T}$, $\Gamma_{j0}^{\xi T}$ and multiplying $\Gamma_{Nh}^{\chi T}$ with $Z_{N}^{T}((Z_{N}^{T})^{\mathrm{T}}\Gamma_{j0}^{T}Z_{N}^{T})^{-1}(Z_{N}^{T})^{\mathrm{T}}$, we obtain matrix K_{N}^{Th}, which in turn we use for projecting from the common factors:

$$\chi_{i,T+h|T}^{NT} = \sum_{j=1}^{N} K_{N,ij}^{Th} x_{jT} \tag{6}$$

where h is the number of step-ahead forecast periods.

As we will show in the following sections, the diversity of our dataset allows for the exploitation of its theoretical advantages, which include the heterogeneity in the fraction of total variance explained by the idiosyncratic components, as well as diversity between the lag structure of the factor loadings, i.e. presence of dynamics. Another advantage of the GDFM is ability to classify of individual variables to leading, coincident and lagging with respect to reference variables at any frequency, most notably the business cycle frequency. The majority, at least of the earlier literature applications included panels of quarterly data. In Marcellino (2006), it is stated that high quality monthly data are a requirement.

3. Data Set, Treatment and Model Parameter Selection

In this section we describe the port dataset, its treatment (deseasonalization outlier removal and standardization), before applying the model and our approach in determining the number of dynamic and static factors, as well as the auxiliary parameters required for implementing the dynamic factor model.

We utilize a panel spanning from January 2005 to March 2012, covering 192 monthly series from 31 U.S. ports and port complexes. These include: Anacortes, Baltimore, Benicia, Canaveral, Crockett, Corpus Christi, Everglades, Hampton Roads, Hueneme, Huston, Kalama, Long Beach, Los Angeles, Miami, Morehead, New York, North Bend / Coos Bay, Oakland, Portland, Redwood City, San Diego, Savannah, Seattle, South Louisiana Stockton, Tacoma, Tampa, Wilmington, and Vancouver. From each port, we utilized all publicly reported series on a monthly basis; these include imported and exported TEUs and / or container-tones, total throughput, dry bulk, general cargo, number of auto units, roll-on roll-off tones, liquid cargo, petroleum, grains, lumber, steel and others, excluding only series containing an extensive number of missing values. Regarding port workforce series, we utilized weekly shifts from several categories for the ports of Long Beach and Los Angeles (combined) and from Hueneme port.

The U.S port dataset described covers up to 84.4% of TEUs and 60.1% of the dry bulk volume of 2012. We have added two proprietary series, namely a) the sum of all available imported TEUs and b) the sum of all available incoming and outgoing container-tones, designated C1. Imp. TEUs and C2. We also included popular indexes related to freight and transportation, such as the Baltic Dry Index (incorporating the demand side of the international dimension of trade), TSI, TSI (freight), TSI (passengers), and the Cass US Freight Export and Shipments indexes. Finally, apart from the GDP, which we converted to monthly frequency by using cubic spline interpolation, we chose to include in the dataset three national economic indexes: Industrial Production (IP), IP: Manufacturing and IP: Materials, in order to evaluate their potential as reference variables. Our first priority has been to maximize the number of series included, sacrificing potential information that could be introduced by a longer panel. We chose to keep the number of external series low, in order a) to limit, to the maximum possible extent, their contribution to the factor model and, b) to reduce the number of series that we would have to obtain temporary forecasts, in the context of a real forecasting exercise. Sources of our dataset include websites of individual U.S. ports, port authorities, Maritime Administration (MarAd) for the cruise passenger series, the Pacific Maritime Association for port workforce shifts and the Saint Louis FED for the economic series, TSI and Cass indexes.

With respect to preliminary data treatment, we removed seasonality and corrected outliers using TRAMO described in Gomez and Maravall (1996), we took log differences for all series to obtain month-to-month growth rates and we mean-standardized the data. As suggested by Altissimo et al. (2001) we did not use SEATS in order to avoid using bilateral filters, since this would introduce revision requirements.

One of the first steps in assessing dynamic factor models is determining the number of dynamic and static factors, as well as auxiliary model parameters. For the former two, we utilized the criteria of Hallin and Liška (2008) and the Alessi et al., (2010) respectively and we cross-checked the results with the heuristic approach of Forni et al., (2000). The results varied between 1 and 3 factors. We also noticed that the majority of the Hallin and Liška (2008) trials pointed to 2, and marginally 3 dynamic factors when we used 'shallow' subpanel runs (from 90% to 95% of the panel size), When a larger number of subpanels were selected, the results tended to point to one factor. Banerjee and Marcellino (2006), recommend the number of factors used in the model should be either equal or larger than the 'true' number, so we chose 2 dynamic factors for our core scenario runs. The choice of the lag window M also affected the number of factors. With $M = round(0.75T^{0.75})$ the information criterion pointed to 2-3 factors, while with $M = round(0.5T^{0.5})$, to one. Nevertheless, the as we will show, the results are quite robust to the number of dynamic factors. For the number of static factors we utilize a range between 7 and 11 factors. In our core scenario, we use 9 static factors, where we obtained the –averaged among forecasting horizons- results. With respect to the number of series included in the panel, we opted for keeping them all, since by removing noisy series with low commonality ratios and high idiosyncratic variance we observed a drop in forecasting performance; a result in line with Nieuwenhuyze (2006). We assume this could be a possible side-effect of disaggregate data: Even though the relevance of a noisy series (e.g. from a small port) to the factor model can be low, its contribution contains unique information exploitable for forecasting. Also, as per Reijer (2005) the distinction between oversampling effects and noisy data, can be difficult.

4. Results, forecasts and discussion

4.1 Panel Dynamics and Model Results

In this section we assess the dynamics of our dataset; we evaluate the commonality of major variables, as well as the lead / lag relationship between port and macroeconomic series.

As suggested by D'Agostino and Giannone (2007), we firstly assess the difference between the number of static and dynamic factors[1] (as calculated by principal components and dynamic principal components respectively), required to capture more than 50% of the panel variance, as an indication of its dynamics and internal lead-lag structure. We then apply the GDFM and review the distribution of the idiosyncratic component variance.

Although the U.S data set in D'Agostino and Giannone (2007) – almost identical to the Stock and Watson (2002) one- is comprised of a large number of aggregate economic variables, real and nominal, asset prices, the yield curve, surveys and other, we report similar findings: The U.S port sector dataset shows enhanced dynamics and a rich lead-lag structure. More than 14 static factors are required to describe 50% of the panel variance, explained by 4 dynamic factors. We consider this result as unexpected; albeit useful for documenting phase-difference information carried by our port trade dataset.

Table 1

Variance explained by principal components

n. of factors	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
static	0.08	0.13	0.18	0.22	0.25	0.29	0.32	0.35	0.38	0.41	0.43	0.45	0.48	0.50	0.52
dynamic	0.22	0.36	0.46	0.56	0.64	0.70	0.76	0.80	0.84	0.87	0.89	0.91	0.92	0.93	0.94
dynamic(b)	0.14	0.24	0.32	0.39	0.45	0.50	0.55	0.59	0.63	0.67	0.70	0.73	0.76	0.79	0.81

Table 2

Distribution of idiosyncratic variance

Variance range	(.0-.1)	(.1-.2)	(.2-.3)	(.3-.4)	(.4-.5)	(.5-.6)	(.6-.7)	(.7-.8)	(.8-.9)	(.9-1)
% of series (2 d. factors)	0.00	0.07	0.05	0.12	0.15	0.17	0.17	0.14	0.13	0.02
% of series (1 d. factor)	0.00	0.01	0.04	0.05	0.11	0.11	0.12	0.14	0.22	0.20

The strong comovement indicated by the low number of factors was expected, since in theory our variables share the same underlying forces. When explained by 2 dynamic factors, the relatively dispersed distribution of the idiosyncratic components in our 'port factor model' (PFM) is similar to the one reported in D'Agostino and Giannone (2007), albeit shifted towards higher idiosyncratic variances, which suggests that the potential benefits of the GDFM apply to the U.S port sector dataset as well. When we use one dynamic factor, the distribution is shifted further to the left, as expected by the disaggregate nature of the dataset.

[1] For reference purposes we provide the estimation of dynamic principal components with $M = round(0.75T^{0.75})$, denoted in Table 1 as dynamic (b).

After applying the GDFM to our panel, using the parameters discussed in the previous section, we compute for each series the commonality ratio, the spectral coherence and the angle of the cross spectral density versus reference series at frequencies zero and at the business cycle frequency: The commonality ratio, or degree of commonality is the amount of variance of the common factor divided by its total variance $c_t = \frac{\chi_t}{X_t}$ measures the information content degree of the original series is included in its common factor, i.e. how well the series is described by its common factor, in the context of the information contained in the panel. This metric should not be confused with a regression coefficient, and ince we are focused in the frequency domain, an appropriate measure of fitness between the common components and the common component of a reference variable across all frequencies is the spectral coherence defined by:

$$C_{\chi_i \chi_{ref}} = \frac{\left| \Sigma_N^{\chi_i \chi_{ref} T} \right|^2}{\Sigma_N^{\chi_i \chi_i T} \Sigma_N^{\chi_{ref} \chi_{ref} T}} \tag{7}$$

for series $i = (1,...N)$ and common component of reference series χ_{ref}.

We use GDP and IP:Man, utilizing them also for the categorization of all common components with respect to their cyclical properties by computing the angle $\phi_i(\theta)$, $\theta \in [-\pi ... \pi]$: As per Forni et al. (2001), we interpret each series as:

pro-cyclical, if $\phi_i(0) = \arg(\Sigma_N^{\chi_i^0 \chi_{ref}^0 T}) = 0$

counter-cyclical, if in phase opposition $\phi_i(0) = \arg(\Sigma_N^{\chi_i^0 \chi_{ref}^0 T}) = \pi$,

We then invert signs for the series in phase opposition and compute the angle $\phi_i(\theta^*)$, using a typical business cycle frequency. We choose to set a three year half cycle $\theta^* = \frac{2\pi}{36}$, and compare $\frac{\phi_i(\theta^*)}{\theta^*} = \frac{\arg(\Sigma_N^{\chi_i^{\theta^*} \chi_{ref}^{\theta^*} T})}{\theta^*}$ with $\frac{\tau}{\theta^*}$, where τ is an empirically set angle, which contains the angles of the 'coincident' series at the $[-\tau, \tau]$ interval. Since $\frac{\tau}{\theta^*}$ is in periods (months), we set this boundary at 3 months. Therefore if $\frac{\phi_i(\theta^*)}{\theta^*} \geq 3$, the series is characterized as leading, if $\frac{\phi_i(\theta^*)}{\theta^*} \leq -3$, lagging, and otherwise as coincident. We report the properties of key individual series as well series major categories in Tables 3 and 4.

Table 3

Commonality Ratios, coherence and cyclical properties of key PFM series

Series	c_t	Rank	$C_{\chi_i \chi_{GDP}}$	ϕ_{GDP}	ϕ_{INDPRO}
Industrial Production: Manufacturing	0.81	1	0.99	-1.0	0.7
PFM Imported TEUs (comp. index C.1)	0.81	2	0.93	-0.1	2.3
Long Beach inbound TEUs	0.79	3	0.93	-1.6	0.6
Virginia total TEUs	0.73	6	0.93	-0.5	2.1
U.S GDP	0.72	9	1.00	0.0	1.8
Los Angeles total TEUs	0.70	11	0.83	1.2	3.9
Virginia export TEUs	0.66	12	0.93	-1.2	1.3
Industrial Production	0.60	16	0.99	-1.8	0.0
Los Angeles - Long Beach longshore shifts	0.60	17	0.96	-0.4	1.8
South Louisiana total dry bulk tones	0.60	19	0.31	9.7	11.4
Savannah TEU throughput	0.59	23	0.96	-0.8	1.6
Baltimore import RoRo tones	0.55	29	0.72	-7.1	-3.9
Oakland container tonnes	0.55	30	0.55	5.1	8.1
Miami cruise passengers	0.52	48	0.63	-0.3	0.6
Cass US Freight Export Index	0.51	50	0.97	-2.1	-0.2
New York export TEUs	0.48	57	0.87	3.0	5.4
Transportation Services Index	0.46	61	0.99	-0.5	1.3
San Diego autos tones	0.37	91	0.96	-1.7	0.0
Corpus Christi liquid bulk tones	0.33	98	0.50	5.1	-7.5
Baltic Dry Index (BDI)	0.29	108	0.09	-7.7	-11.0
Tacoma assessable containers	0.27	117	0.81	5.9	8.5
Los Angeles general cargo tons	0.14	165	0.96	-1.9	-0.7

Table 4

Summary properties of PFM series

Series Group	Number of series	c_t	$C_{\chi_i \chi_{GDP}}$	$C_{\chi_i \chi_{IP:MAN}}$	$\phi_{IP:MAN}$
Total TEUs (throughput)	10	0.49	0.78	0.78	2.5
Imported TEU / TEU-tonnes	7	0.54	0.92	0.93	-0.1
Exported TEU / TEU-tonnes	10	0.44	0.81	0.77	3.6
Port workforce shifts	5	0.35	0.87	0.89	0.9
Liquid Bulk tonnes	6	0.37	0.63	0.68	0.5
Cruise Passenger series	9	0.27	0.67	0.67	1.8
Empty TEU / TEU-tonnes	16	0.27	0.58	0.63	1.8
Auto units / RoRo-tonnes	20	0.24	0.66	0.71	0.7
Bulk tonnes	33	0.25	0.47	0.49	4.9
General Cargo tonnes	15	0.18	0.67	0.67	1.8
Breakbulk tonnes	3	0.12	0.32	0.34	-7.1

We point out the high commonality ratios c_t of U.S GDP and IP: Manufacturing and we consider this result unexpected, with respect to the level of disaggregation for the majority of the series included in the panel. This means that the 2 dynamic factors of the PFM reflect the same driving forces of the U.S. economy and manufacturing. The results are consistent also using one common factor, with the commonality ratio of GDP at 0.70. Even if we exclude nearly all aggregate series from our dataset (TSI(Passengers),TSI, IP:Mat and the 2 Cass indexes) the commonality ratio of the U.S GDP remains high at 0.71 with two dynamic factors). The ranking of series groups based on their commonality ratio averages is dominated by series relative to TEUs and container-tones, followed by port workforce series. Bulk, general cargo and breakbulk present the lowest commonality ratios, reflecting the different effect they have on the panel. However relevant to the GDP, we can see the elevated importance of the auto / RoRo series PSD and the comovement of general cargo tones with IP: Manufacturing. The highest commonality ratios in the first quartile are dominated by series from the large ports of Los Angeles, Long Beach, port of Virginia, Huston and South Louisiana, most of which involve imported TEUs / container tons and container totals, or throughput, including incoming, outgoing and empty containers. Exceptions to this finding are the Ports of Virginia exported TEUs, liquid and dry bulk totals, port of Huston exported TEUs and clerk and longshore shifts for ports of Los Angeles and Long Beach combined.

Table 5

Cyclical Properties of PFM series with respect to GDP

Series	Leading	Coincident	Lagging	Total
Pro-cyclical	44	75	49	168
Counter-cyclical	4	3	17	24
Total	48	78	66	192

The cyclical properties of our port dataset are relatively balanced, although the average lead and lag are lower compared with macroeconomic datasets. Major leading series (on average 4-7 months) include exported TEUs from New York, Baltimore, Huston and Savannah, TEU / container tones throughput from Tacoma, Los Angeles, Oakland, Portland, cruise passengers total and BDI. The highest leads (13-17 months) are found in Huston export empty TEUs, South Louisiana throughput and Vancouver auto tones. The highest lags, ranging from 11 to 15 months are found in Morehead breakbulk tones, Baltimore steel and other metals tons and Oakland general cargo tones. Most of the 24 counter-cyclical series refer to bulk cargo tones and empty containers. Exceptions are the lagging countercyclical series of Willimington TEUs and Oakland autos. All aggregate indexes and most port workforce series are found to be coincident with the GDP and Industrial Production. It should be noted that the categories of general cargo, bulk and breakbulk are less consistent with respect to their cyclical properties and more sensitive to model parameter changes.

Our results, most evident in larger ports, are analogous with the ones presented in Altissimo et al., (2001) where the commonality ratios for all sectors of the Euro Area were assessed between 0.50 and 0.60, with the trade sector being on the high end (0.58), but less correlated than other variables with the business cycle. Imports in Altissimo et al., (2001) were found to be lagging, compared with exports, although some commodities such as raw materials, crude oil and beverages were found to lead the cycle. Our results with respect with oil are mixed: South Louisiana crude oil and Corpus Christi oil series are lagging and lagging/coincident respectively, while South Louisiana petro-chemicals and total liquid bulk series lead by 10.5 months on average. Nieuwenhuyze (2006) also reports a similar relationship between the common factors of Belgium's imports and exports, to our results between imported and exported TEUs or TEU tones: Commonality ratios 32.6 and 40.9 for exports and imports respectively, with imports slightly lagging versus exports. Two additional observations can be deduced from table 4: a) the shift in raking between commonality ratios and spectral coherence vs. GDP for port workforce shifts, general cargo tones and auto units / RoRo tones, indicating that their relatively lower explanatory power for the factor model does not imply that they are less 'appropriate' than other series categories for GDP regressions. b) The PFM, through the common component of the GDP provides strong indications that it contains different, or additional, non-idiosyncratic information content for the U.S economy compared with popular national indexes related to freight, such as the TSI (Freight) and two Cass indexes. This information can be possibly related to the amount of incoming / imported freight or materials consumed or processed in proximity to the port hinterland / region, as well to the port added value. Identifying these differences is an item for further work.

4.2 Design and Evaluation of Port Factor Model Forecasts

Due to the length of our series (87 months) we opt for using an expanding (also called recursive) window out-of-sample forecasting approach: We treat the first half of the panel $(T/2)$ as known and we compute forecasts of all series for horizon h, then extending the panel to $(T/2)+h$, iterating up to $T+h$. Cheung and Demers (2007) report that the best forecasting properties are obtained by their smallest rolling window sample, set at 40 quarterly observations. For our forecasting exercise we compute the relative mean square forecast errors (RMSFE) for each of the series, using a) the PFM b) a naïve model, in which the 'forecast' is the value of the previous period, and c) a random walk with drift, constrained within its past variance.

Table 6

US. GDP Relative RMSFEs with Naïve, Random Walk and Port Factor Model

Periods ahead	PFM	Naive	Random walk	DM Test Naïve vs. PFM(lev.)	DM Test RW vs. PFM(lev.)	DM Test Naive vs. PFM(g.r)	DM Test RW vs. PFM(g.r)
1	2.33	3.63	3.19	-2.42**	-1.49	2.65	-1.47
2	2.04	7.16	3.17	-2.59***	-2.27**	2.02	-1.38
3	1.62	10.39	2.91	-2.58***	-2.27**	0.89	-1.76*
4	1.31	13.20	3.10	-2.51**	-2.47**	-0.96	-1.92*
5	1.21	15.54	3.02	-2.43**	-2.48**	-1.79*	-2.25**
6	1.08	17.49	3.12	-2.39**	-2.53**	-2.08**	-2.39**
7	1.00	19.15	3.17	-2.44**	-2.68***	-2.48**	-2.41**

We designate *,**,*** as 10%, 5% and 1% significance levels respectively, in which equal predictive ability between rival models and GDFM is rejected

We report the RMSFE and the Diebold-Mariano (1996) test results (DM) for the consistency of forecast gains, for the U.S GDP, IP and a sample port-specific series, the port of Long Beach outbound TEUs both for the series levels (lev.) and for growth rates (g.r). In terms of RMSFEs, the GDFM outperforms both the naïve and the random walk level forecasts in all forecasting horizons. The DM tests for level forecasts reject in most cases the hypothesis for equal forecasting accuracy between GDFM and rival forecasts. In the case of GDP growth rates, the DM tests fail for horizons 1-3. This is not surprising, since the 'monthly' GDP values are a product of interpolation; in a sense, the GDP unobservable monthly growth rates are constructed by a naïve model. Apart from this exception, GDFM forecasts outperform both rival methods.

GDFM forecasts improve, as the horizon increases, which provide additional incentives for compiling longer datasets[2.] However, for the series of Long Beach the rate of improvement over forecasting horizons in terms of RMSFE is more gradual, visible also in the DM level forecast results. The lower DM tests results of the growth rates compared with level forecasts

[2] The maximum number of 'periods ahead' forecasts depend on the number of lead / lag Bartlett window, which in turn is defined as a function of the number of periods T covered by the panel.

was expected, since the former is a transformation or the latter, producing MSFEs several orders of magnitude lower than the ones produced by levels. Having presented the results for three series with high commonality ratios, we summarize the DM tests for all 192 series and horizons (1-7) for key confidence levels. The majority of the DM results are globally shifted in favor of the PFM, both vs. the naïve and the RW method, with very few exceptions having positive values. Even at the challenging environment of growth rate forecasts against the random walk, in more than 20% of the instances a 5% confidence level is achieved, and in the case of level forecasts, in more than 66% of the instances. We also report that the highest results for the PFM are obtained in horizons 4 to 7 months. We consider these results as promising, taking into account that our evaluation timeline spans within a volatile period, influenced by the 2008 shock. The lowest results in DM tests are concentrated in series related to bulk cargo, and smaller ports, such as Aberdeen, Morehead, Vancouver, as well as the BDI. Low DM results in container related series include the Los Angeles and Virginia (Hampton Roads) incoming TEUs.

Table 7

IP index Relative RMSFEs with Naïve, Random Walk and Port Factor Model

Periods ahead	PFM	Naive	Random walk	DM Test Naïve vs. PFM(lev.)	DM Test RW vs. PFM(lev.)	DM Test Naive vs. PFM(g.r)	DM Test RW vs. PFM(g.r)
1	2.36	6.02	4.64	-1.22	-2.43**	-0.88	-1.55
2	2.22	9.04	5.08	-1.91*	-2.35**	-1.30	-1.41
3	2.23	12.80	4.66	-2.14**	-2.43**	0.03	-0.95
4	2.39	16.59	4.91	-2.23**	-2.17**	-1.34	-1.60
5	1.35	19.77	4.49	-2.03**	-2.26**	-1.58	-2.08**
6	1.06	22.25	4.93	-2.12**	-2.60***	-1.95*	-2.24**
7	1.00	24.57	4.64	-1.97**	-2.45**	-1.83*	-2.38**

Table 8

Long Beach Outbound TEUs MSFEs with Naïve, Random Walk and Port Factor Model

Periods ahead	PFM	Naive	Random walk	DM Test Naïve vs. PFM(lev.)	DM Test RW vs. PFM(lev.)	DM Test Naive vs. PFM(g.r)	DM Test RW vs. PFM(g.r)
1	1.81	4.19	4.03	-3.66***	-2.30**	-2.94***	-1.86*
2	1.72	4.76	3.81	-1.71*	-2.28**	-0.55	-1.28
3	1.54	6.02	3.55	-2.43**	-2.42**	-1.57	-1.74*
4	1.62	7.46	3.80	-2.16**	-2.64***	-2.77***	-2.27**
5	1.29	7.96	3.76	-1.75*	-2.60***	-1.44	-1.60
6	1.27	8.53	4.05	-1.83*	-2.64***	-2.45**	-2.04**
7	1.00	9.06	4.05	-1.53	-2.35**	-2.57**	-1.61

Table 9

DM tests distributions for level and growth rate forecasts

	Level forecasts				Growth rate forecasts			
DM test range	PFM vs. Naïve Freq.	PFM vs. Naïve Cum.%	PFM vs. RW Freq.	PFM vs. RW Cum.%	PFM vs. Naïve Freq.	PFM vs. Naïve Cum.%	PFM vs. RW Freq.	PFM vs. RW Cum.%
-2,58***	452	33.63%	201	14.96%	146	10.86%	10	0.74%
-1,96**	441	66.44%	708	67.63%	336	35.86%	285	21.95%
-1,65*	213	82.29%	202	82.66%	221	52.31%	335	46.88%
-1,28	168	94.79%	205	97.92%	249	70.83%	327	71.21%
0	51	98.59%	27	99.93%	328	95.24%	369	98.66%
More	19	100.00%	1	100.00%	64	100.00%	18	100.00%

On the other hand, for the outgoing TEUs of Los Angeles and Long Beach, exported and imported TEUs of Baltimore and exported TEUs of Savannah and Huston, as well as for the Cass and TSI indices, the PFM has produced competitive results. We also obtained good results for several of the series related with vehicles (auto units, RoRo tons etc.). The latter is supported by Kitchen and Monaco (2003), where Total Light Vehicle Sales Growth series was assessed as one of the best scoring relative RMSE series in short (1-3 months) forecasting utilized in the US Treasury Real-Time Forecasting System (RTFS), a methodological approach similar to the factor model.

Our parameter estimation strategy in this exercise was to establish a consistent basis of dynamic, static factor ranges and Bartlett lag window size with respect to maximizing the *average* forecasting performance across all variables. We realize that this strategy may produce inferior forecasts of some variables, since the optimal model parameter setup for forecasting varies both between series and forecast horizons, and more importantly, non-linearly to the number of static factors. Therefore, we assume that there is room for improvement to the forecast accuracy of the model, for several series included in the panel. Not all series share the same dynamics, or that one model specification is optimal for all series. Customization of the PFM settings for specific ports, cargo types or forecast horizon improves their forecasts, at the expense of worse forecasts for other series.

5. Conclusions and Future Work

In this exercise, using a large dataset from the majority of U.S ports, covering 84.4% of TEUs and 60.1% of the dry bulk monthly volumes, as well as cruise passengers between 2005-2012 in the context of a one-sided generalized factor model we have a) assessed its dynamics and rich lead-lag structure reflected in 2 dynamic factors, b) demonstrated favorable forecasting performance for series of individual ports as well as for aggregate economic indices, especially in longer forecast horizons 4-7 months, c) we assessed high commonality ratios of the U.S GDP and IP indices common factors, reflecting the link between ports, trade and a national economy, d) we pointed that the common component of the GDP, as assessed by our 'port factor model' provides strong indications that it contains different or additional non-idiosyncratic information content for the U.S economy compared with popular freight indexes, such as the TSI or the Cass indices. The results of this approach provide added value in two areas of interest in contemporary research: Firstly in port economics, by providing competitive forecasts for individual incoming or outgoing cargo types, as well as port throughput and secondly, in macroeconomic applications of national factor models.

The results provided, are useful both for the U.S port industry, as well as for government authorities, agencies and central banks: Short and mid-term forecasting is of interest to port operators and authorities, in order to prepare or adapt to future incoming or outgoing trade volume fluctuations by adjusting labor requirements, outsourcing decisions, delay or rush infrastructure and equipment maintenance, as well as enhance operational coordination with freight transport operators. Our proposed approach offers potential advantages, due to the fact that forecasts are constructed based on contemporary and past information not only for the port series of interest, but from all ports at once. From a government agency or a central bank perspective, the links of ports and the national economy can contribute as an additional asset for macroeconomic forecasting, as well as creation of timely available port and trade related indexes.

With the compilation of larger datasets, apart from the obvious extension of the forecasting horizon, and by utilizing a rolling window approach we will be able to explore potential bias-variance trade-offs and infer possible heterogenic processes or structural changes over time, as well as evaluate alternative factor model implementations. The next logical step is the construction of a coincident indicator, utilizing as basis the common factor of the U.S GDP, as explained by our model. Furthermore the U.S port sector, as mapped in this exercise can provide a constrained environment which may shepherd the identification of demand or supply shocks within the assessed common factors.

Acknowledgements

This work was partially funded by Transport for Athens. We would also like to thank M. Brooks and A. Deloukas for their comments and I. Minis, E. Valavani and E. Pilafidis for inspiration.

References

Al-Hassan, A., 2009. A coincident indicator of the Gulf Cooperation Council (GCC) business cycle. International Monetary Fund.

Ajevskis, V., and Dāvidsons, G., 2008. Dynamic Factor Models in Forecasting Latvia's Gross Domestic Product. Latvijas Banka,Working Paper No. 2008/02.

Alessi, L., Barigozzi, M., and Capasso, M., 2010. Improved penalization for determining the number of factors in approximate factor models. Statistics and Probability Letters, 80(23), 1806-1813.

Altissimo, F., Bassanetti, A., Cristadoro, R., Forni, M., Lippi, M., and Veronese, G., 2001. EuroCOIN: a real time coincident indicator of the euro area business cycle (Vol. 3108). CEPR discussion paper.

Aruoba, S. B., Diebold, F. X., Kose, M. A., and Terrones, M. E., 2010, September. Globalization, the business cycle, and macroeconomic monitoring. In NBER International Seminar on Macroeconomics 2010 (pp. 245-286). University of Chicago Press.

Banerjee, A., and Marcellino, M., 2006. Are there any reliable leading indicators for US inflation and GDP growth?. International Journal of Forecasting, 22(1), 137-151.

Barhoumi, K., Darné, O., and Ferrara, L., 2010.. Are disaggregate data useful for factor analysis in forecasting French GDP? Journal of Forecasting, 29(1-2), 132-144.

Baxter, M., and Kouparitsas, M. A., 2005. Determinants of business cycle comovement: a robust analysis. Journal of Monetary Economics, 52(1), 113-157.

Brillinger, D. R., 1981. Time Series Data Analysis and Theory, New York: Holt, Rinehart and Winston.

Carriero, A. and Marcellino, M., 2007. A comparison of methods for the construction of composite coincident and leading indexes for the UK. International Journal of Forecasting, 23(2), 219-236.

Cheung, C. and Demers, F., 2007. Evaluating forecasts from factor models for Canadian GDP growth and core inflation (No. 2007, 8). Bank of Canada Working Paper.

Chlomoudis, C. I. and Pallis, A. A., 1996. Investment policies in ports' infrastructure in the perspective of the European Shortsea Shipping Networks: The case of Greece. Proc. ESSS, 96.

Chlomoudis, C. I. and Pallis, A. A., 1997. Investment In Transport Infrastructure: Have the EU initiatives promoted their balanced and rational distribution? World Transport Policy and Practice.

D'Agostino, A. and Domenico G., 2007. Comparing alternative predictors based on large-panel factor models. CEPR Discussion Papers 6564.

Doz, C., Giannone, D. and Reichlin, L., 2011. A two-step estimator for large approximate dynamic factor models based on Kalman filtering. Journal of Econometrics, 164(1), 188-205.

Diebold, F. X., and Mariano, R. S., 2002. Comparing predictive accuracy. Journal of Business and economic statistics, 20(1).

Forni, M., Hallin, M., Lippi, M. and Reichlin, L., 2000. The generalized dynamic-factor model: Identification and estimation. Review of Economics and statistics, 82(4), 540-554.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L., 2001. Coincident and leading indicators for the euro area. The Economic Journal, 111(471), 62-85.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L., 2005. The generalized dynamic factor model. Journal of the American Statistical Association, 100(471).

Gomez, V. and Maravall, A., 1996. Programs TRAMO (Time series regression with arima noise, missing observations, and outliers) and SEATS (Signal extraction in arima time series). Instructions for the User. Documento de Trabajo, 9628.

Guichard, S., and Rusticelli, E., 2011. A dynamic factor model for world trade growth. OECD Economics Department Working Papers No. 874

Hallin, M. and Liška, R., 2007. Determining the number of factors in the general dynamic factor model. Journal of the American Statistical Association, 102(478), 603-617.

Kapetanios, G. and Marcellino, M., 2006. A Parametric Estimation Method for Dynamic Factor Models of Large Dimensions (No. 5620). CEPR Discussion Papers.

Kitchen, J. and Monaco R., 2003. Real-time forecasting in practice: The US Treasury staff's real-time GDP forecast system, 10-19.

Lahiri, K. and Yao, W., 2004. A dynamic factor model of the coincident indicators for the US transportation sector. Applied economics letters, 11(10), 595-600.

Lahiri, K. (2010). Transportation Indicators and Business Cycles (Vol. 289). Emerald Group Publishing.

Marcellino, M., 2006. Leading indicators. Handbook of Economic Forecasting, 1, 879-960.

National Cooperative Freight Research Program (2014). Performance Measures for Freight Transportation (NCFRP Report 10), Washington, DC: Transportation Research Board, p. 126

Nguiffo-Boyom, M., 2008. A monthly indicator of Economic activity for Luxembourg (No. 31). Central Bank of Luxembourg.

Nieuwenhuyze, C. V., 2006. A generalised dynamic factor model for the belgian economy-useful business cycle indicators and GDP growth forecasts. National Bank of Belgium Working Paper, (80).

Perevalov, N., and Maier, P., 2010. On the advantages of disaggregated data: Insights from forecasting the US economy in a data-rich environment (No. 2010, 10). Bank of Canada Working Paper.

Reijer, A. D., 2005. Forecasting Dutch GDP using large scale factor models. Netherlands Central Bank, Research Department.

Reijer, A. D., 2007. Identifying Regional and Sectoral Dynamics of the Dutch Staffing Labour Cycle. Netherlands Central Bank, Research Department.

Schumacher, C. and Dreger, C., 2002. Estimating large-scale factor models for economic activity in Germany: Do they outperform simpler models?

Schumacher, C., 2007. Forecasting German GDP using alternative factor models based on large datasets. Journal of Forecasting, 26(4), 271-302.

Stock, J. and Watson, M., 2002a. Forecasting using principal components from a large number of time predictors, Journal of the American Statistical Association 97, 1167-1179.

Stock, J. H. and Watson, M.W., 2011. Dynamic factor models. Oxford Handbook of Economic Forecasting. Oxford University Press, USA, 35-59.

Tsamourgelis, I., Paflioti, P., and Vitsounis, T., 2013. Seaports Activity (A)synchronicity, Trade Intensity and Business Cycle Convergence: A Panel Data Analysis. International Journal of Maritime, Trade and Economic Issues (IJMTEI), 1(1), 67-92.