A dynamic log-linear regression model to forecast numbers of future filings at the European Patent Office

Peter Hingleya,*, Walter Parkb

a European Patent Office, Munich, Germany
b American University, Washington, DC, USA

ABSTRACT

An econometric model is applied to forecast future levels of patent filings at the European Patent Office out to 2019, using historical data from 1990 to 2013 with 28 source country terms. Descriptors include Research and Development expenditures and Gross domestic product, where the latter is split into trend and business cycles components. The model is applied to logarithmically standardised data.

The effects on the forecasts of additional future positive and negative stimuli to the GDP components are considered. Reasonable forecasting accuracy is found. Using a series of shorter historical data windows may give improved accuracy for short term forecasts.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The European Patent Office (EPO) forecasts future patent filings in order to plan for likely workloads in the patent granting process, such as expected numbers of searches, substantive examinations, grants and renewals. These plans have implications for the requirements for staff and infrastructure. Procedurally, there is an annual cycle that proceeds from the forecasts via a business plan to finalisation in a budget document [1]. This budget is renewed annually and covers five years beyond the year in which it is produced.

The time series that are to be forecasted are shown in Fig. 1 with data up to 2013. EPO filings are a mixture of different types. Here we will consider forecasting the sum of Euro-direct filings and Euro-PCT international phase filings (Total filings in Fig. 1), after removing divisional filings (a form of retrospective Euro-direct filing that is forecasted separately). Other types of filings and downstream workload forecasts are then usually obtained by applying ratios to the forecasts for Total filings.

A variety of approaches are available that are based on historical data [1–3] or surveys [4]. The regression method that will be considered involves a dynamic log-linear model for annualised data that has been used since 2007 [5]. This operates on transformed EPO Total filings from 28 countries or regions, with autoregressive terms as well as source country Gross Domestic Product (GDP) and Research and Development expenditures (R&D) as independent variables. The model has recently been extended to consider the effects of business cycles [6]. This paper discusses the way that the approach has been customised for the forecasting process at EPO.

Matters of particular concern include transforming the data to achieve stationarity, how to calculate confidence intervals for the filings forecasts and how to interpret the forecasts and their accuracy against the later outcomes. The paper is organised as follows. Section 2 explains the model. In Section 3.1 a panel data set from 1990 to 2013 is fitted both to a model in levels and to a model in year-to-year differences. Section 3.2 shows the forecasts and interprets them. Section 4 considers the effect of a hypothetical boom or recession for one year during the forecasted period and also a scenario that is based on assumptions about the shape of the future business cycle. Section 5 looks at stability by fitting subsets of the same data in terms of a number of overlapping time windows. Section 6 discusses further directions.

2. The model for making parameter estimates and forecasts

The following regression model is used for EPO Total filings from a source country:-
Total filings $P$ are transformed as indicated to $\log(P/L)$. This allows for a standardisation between countries, as $L$ is treated as a proxy for country size, and for stabilising error by the logarithmic transformation. Based on [10], the value of $R$ is lagged by five years in order to incorporate the concept that R&D expenditures have their effect after a delay. Most EPO filings are subsequent filings that take place up to a year after first filings, so the assumption is that R&D expenditures “cause” first filings about 4 years later on. Qualitatively similar results are obtained via a model with no lag in $R$ [6], and in Section 5 below some comparisons are made between lags of 1, 3, 5 and 7 years. No time dummies are included, which gives better forecasting ability by assuming that the process remains stable over time.

The GDP term $Y$ is decomposed via the Hodrick and Prescott filtering method [11] into its trend and cyclical components ($Y^T$ and $Y^c$ respectively) and then the business cycle variable is $u = Y^c/Y^T$. This is detailed in Ref. [6], where it is demonstrated that the usage of $u$ and $Y^c$ rather than $Y$ improves the goodness of fit to the model for filings on the historical training data set, and so may also provide improved forecasting ability.

Annex 1 indicates the way that the forecasts for EPO filings from the source countries and their variabilities were calculated and combined to make the forecasts for Total filings. The authors will be prepared to share further details of the methods on request.

### 3. Results

#### 3.1. Fitting the models

The analysis here reflects the data up to 2013 that were available in the second half of 2014. The model is fitted to a 28 source country-of-origin data set using annualised EPO Total filings from 1990 to 2013. Data for the variables are calculated both as levels and as growth rates.
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات