

Forecasting Patent Applications at the European Patent Office: A Bottom-Up Versus Top-Down Approach

Prepared for

WIPO-OECD Workshop on Statistics in the Patent Field,
11-12 October 2004, Geneva, Switzerland

By

Frederick L. Joutz
Research Program on Forecasting
Department of Economics
The George Washington University
Washington, DC 20052
bjmark@gwu.edu



Acknowledgements: This presentation benefited from helpful comments and suggestions by Peter Hingley, Marc Nicolas, and Costas Mastrogiannis. Any errors or omissions are my own. All opinions are mine and independent of the USPTO.

Benchmark Forecasts

Overview

- This paper presents preliminary results on forecasting patent applications at the European Patent Office using annual data.
- A two step framework is used in the modeling.
 - First Filings
 - Secondary or Subsequent filings
- Two Models are developed.
 - An aggregate model
 - A disaggregate regional model (Europe, Japan, US, and Other)



Benchmark Forecasts

Overview

- using total applications is developed as a function of previous filings and measures of R&D activity. A second model is developed distinguishing between countries/regions of origin and first and subsequent filings. Four regions will be considered: Europe, Japan, United States, and rest of world. First filings by region will depend on perceptions regarding the size of the intellectual property market in Europe, past filings, and economic activity. Subsequent filings are assumed to depend on previous domestic filings in a region, measures of R&D, and perceptions of the intellectual property market in Europe.



Benchmark Forecasts

Overview

- Forecasting patent filings is one of the important issues of the Trilateral Statistical Working Group, WIPO, and the OECD.
- TSWG, composed of the EPO, JPO, and USPTO. The three offices meet at least once a year to discuss this issue among a host of other patenting issues.
- They have been holding annual meetings since 1992.
- The members treat forecasting as an important exercise for planning future resource and manpower requirements and revenues.
- This paper builds on previous research among the TSWG participants and a recent paper by Hingley and Nicolas (2004).



Benchmark Forecasts

A Theoretical Model of Patent Application Filing

- Patents protect more than just intellectual property; they are an intrinsic component of the larger economic picture. This occurs through the process of innovation, technological and scientific change and economic productivity and growth. The process is the result of the demand for and production of “new” knowledge.
- Schmookler (1954) - industrial invention is economically caused.
- In his view invention is driven by the interaction of supply and demand forces.
- Scherer (1983) Pavitt (1982), Hall, Griliches, and Hausman (1986) relationship between R&D effort and patent activities although primarily at the firm level.
- Griliches (1989).
- Adams, Kim, Joutz, Trost, and Mastrogianis (1997)
- Eaton and Kortum (1996 and 1998) and Gardner and Joutz (1996)



Benchmark Forecasts

A Theoretical Model of Patent Application Filing

- The most recent advancement of the endogenous growth theory has been the emergence of R&D-based models of growth in the seminal papers of Romer (1990), Grossman and Helpman (1991a, 1991b) and Aghion and Howitt (1992).
- This class of models agrees with the neoclassical Solow model that capital broadly defined is subject to diminishing returns, and hence the accumulation of capital does not sustain growth in the long run.
- Instead, technological progress is the source of sustained long run growth in both types of models.
- The point of departure lies in the way technological progress is viewed. In the neoclassical model, technology evolves exogenously.
- R&D- based models, the evolution of technology is explicitly and formally modeled as an endogenous process. Technological progress occurs as profit-maximizing firms invest in advanced technologies, and is promoted by the allocation of more productive resources towards R&D.



Benchmark Forecasts

A Theoretical Model of Patent Application Filing

- The model involves four variables: Output (Y), capital (K), labor (L), and technology or knowledge (A). [There are two sectors: a goods-producing sector where output is produced, and an R&D sector where additions to the stock of knowledge are made. Labor can be freely allocated to either of the two sectors, to produce output (LY) or to produce new knowledge (LA). Hence, the economy is subject to the following resource constraint $LY + LA = L$, where L represents the total amount of labor in the economy.
- Specifically, output is produced according to the following production function:

$$(1) \quad Y = K^a (AL_Y)^{1-a}$$



Benchmark Forecasts

A Theoretical Model of Patent Application Filing

The production function approach to knowledge in is the underpinning of the long-term modeling framework

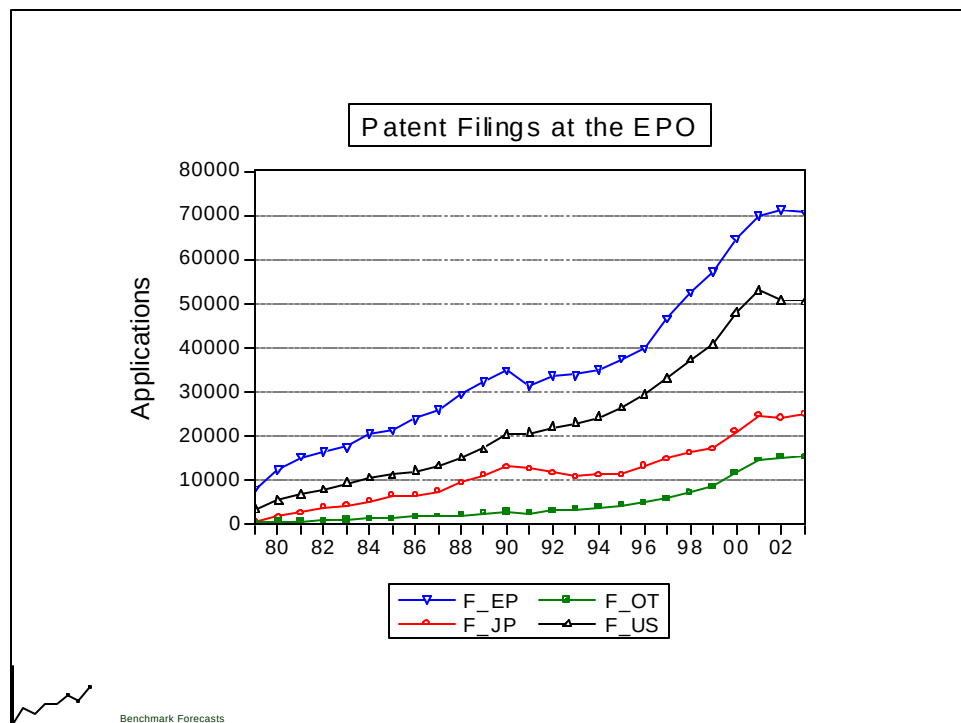
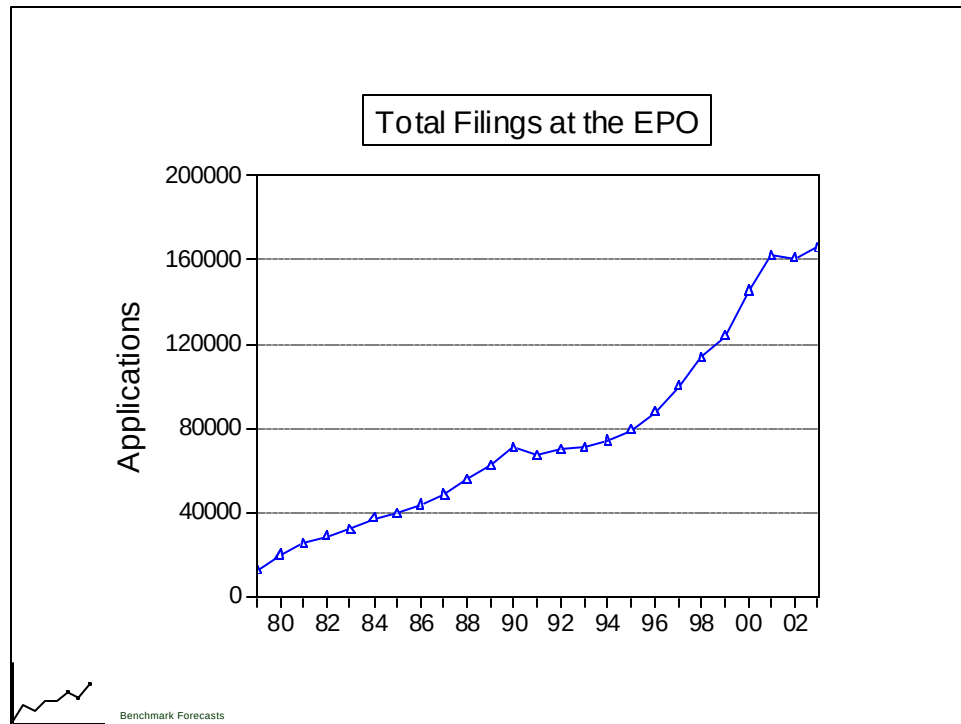
Research labor input is replaced by R&D expenditures as a measure of research effort primarily for data reasons.

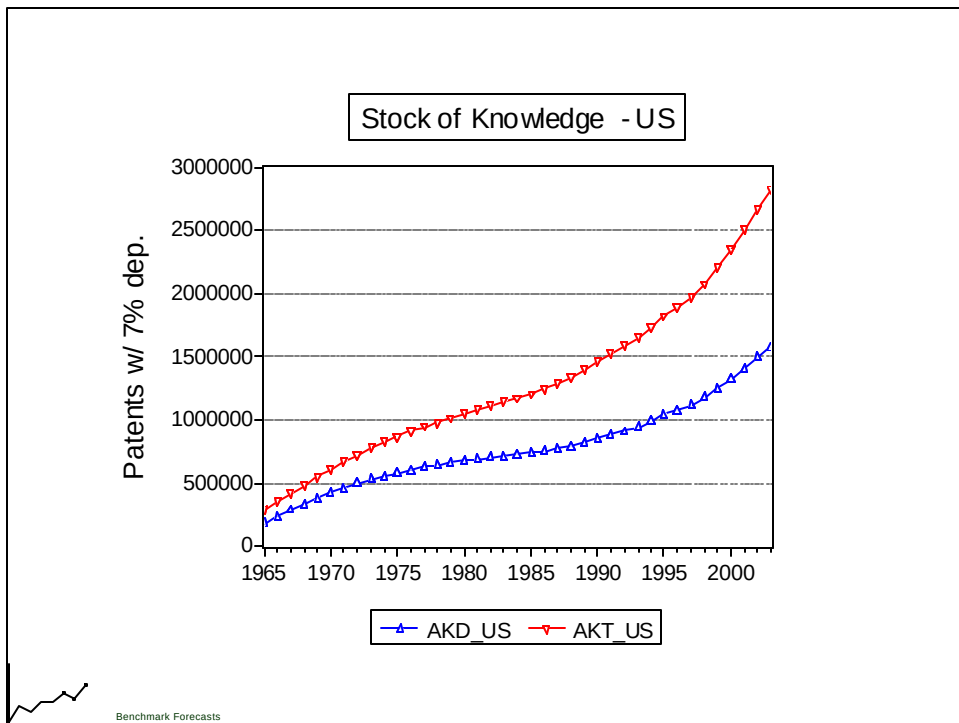
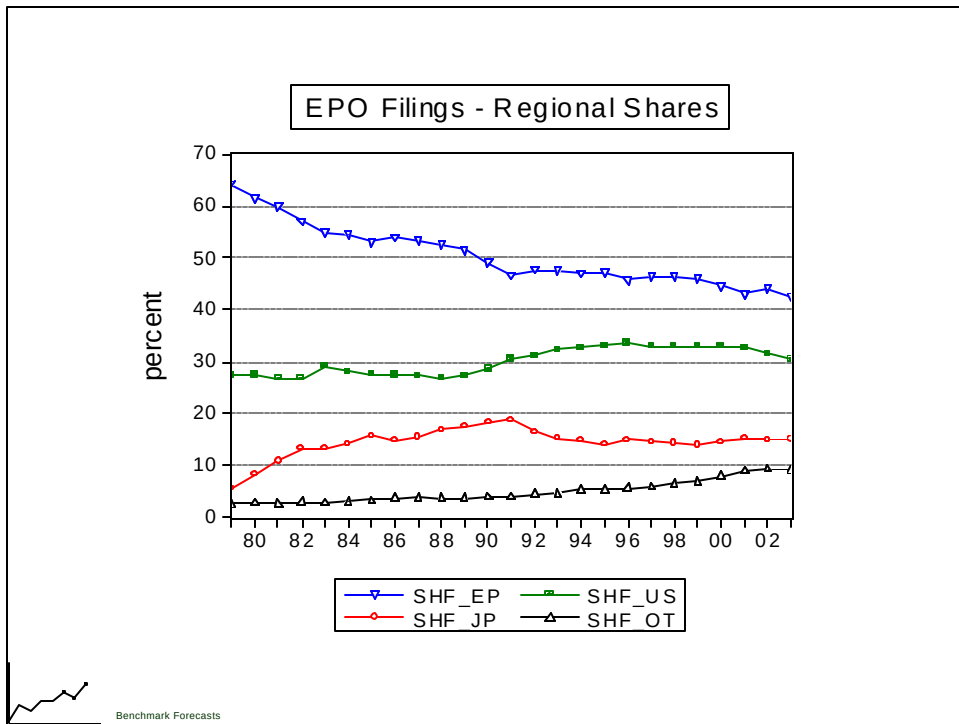
The production function concept is used in a long-term context for generation of new knowledge. is represented by patent application filings and the level of is calculated as the stock of historical patents

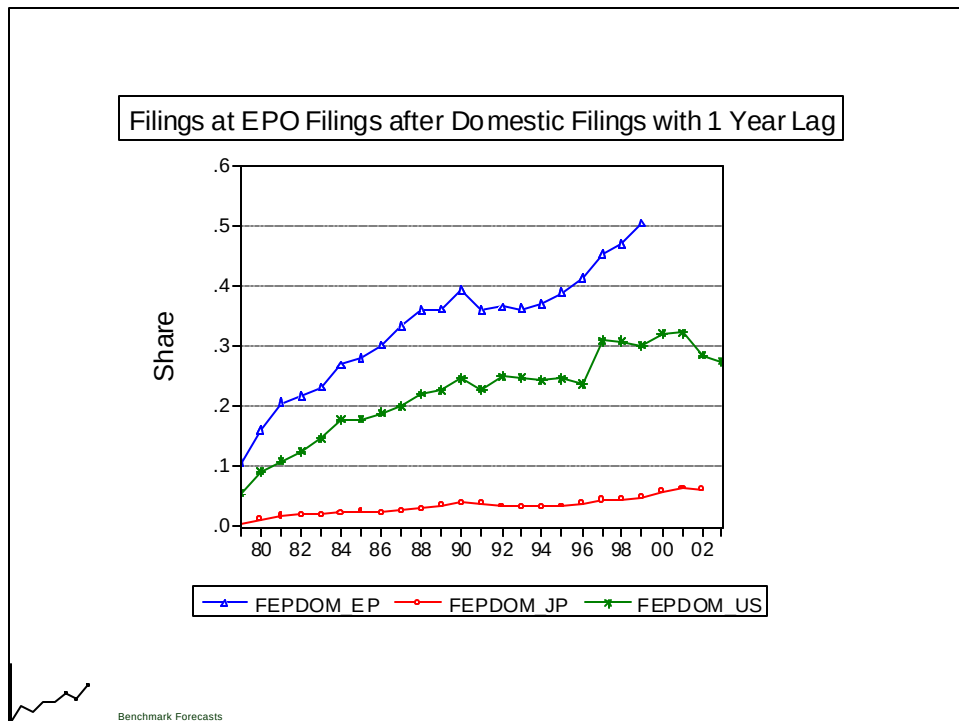
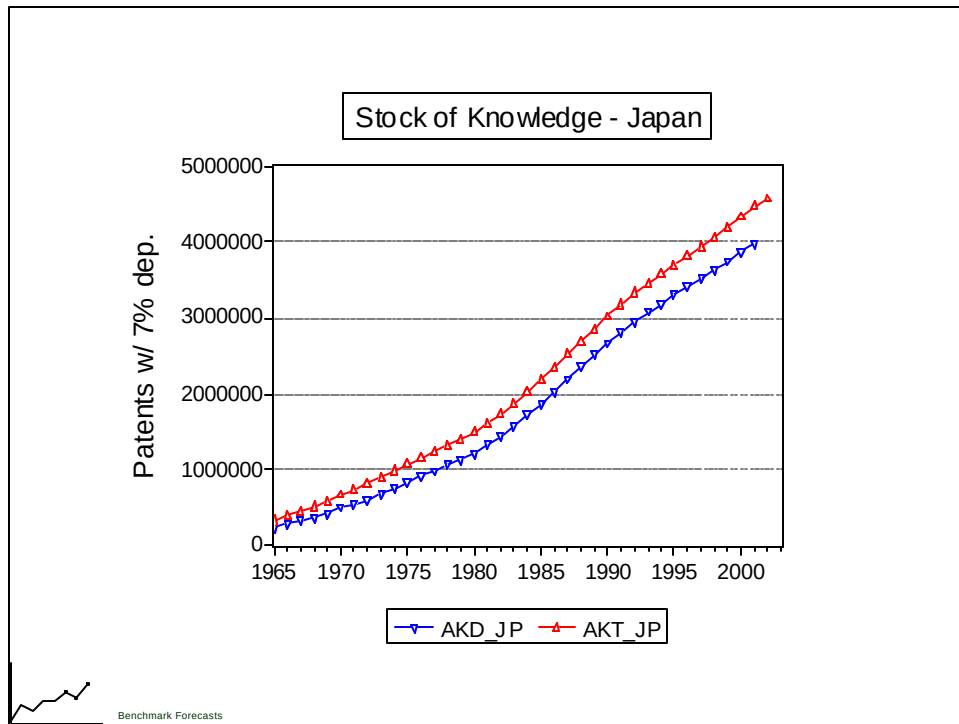
$$\dot{A}_t = d A_t^f L_A^l \approx d A_t^f RD_A^l$$



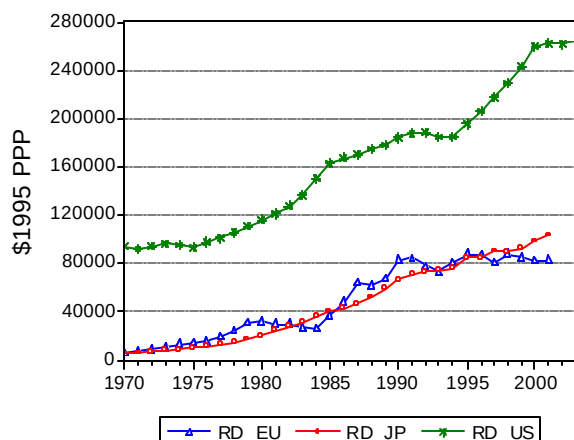
Benchmark Forecasts







Research and Development Expenditures - Europe, Japan, and US



Benchmark Forecasts

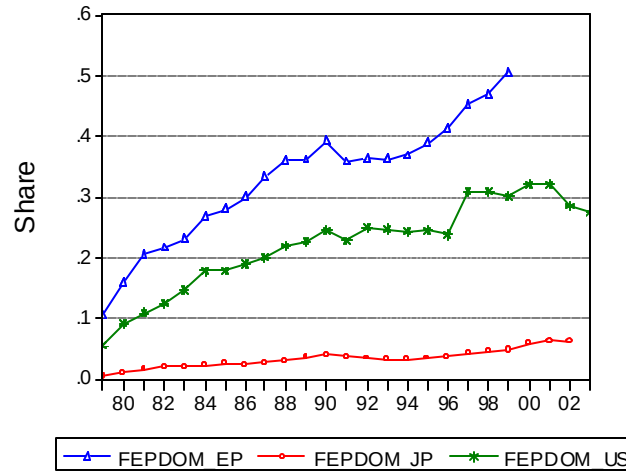
The Modeling Procedure

- Inventors typically first file a patent application in their home country. The first filing represents an indicator of innovative activity.
- Patent protection on an international scale, perhaps based on preliminary searches, is sought about a year later. The preferred route is through an international or supranational procedure to reduce the duplication costs. Currently the European Patent Organization has 31 contracting member countries. This route is referred to as a subsequent or secondary filing.
- The forecasting problem is complicated by the fact that there are multiple routes for patent protection applications. Inventors have the option of filing nationally, through the European system, and the International PCT route administered through the World Intellectual Property Organization.
- However, the primary work load of searches and preliminary examinations from the PCT applications is performed through nine patent offices. The EPO is one of the most important offices authorized or designated to perform this work. This has become increasingly popular as over 90% of the European contracting states are selected when using the PCT route.



Benchmark Forecasts

Filings at EPO Filings after Domestic Filings with 1 Year Lag



Benchmark Forecasts

The Modeling Procedure

- The model framework proceeds in two stages. See Hingley and Nicolas (2004) for a further exposition of this framework.
- In the first stage non-EPO and EPC patent applications are filed domestically.
- The model specification is ADL(p,p), autoregressive distributed lag model based on economic growth theory and the knowledge production function.

$$DomFil_t = b_0 + \sum_{i=1}^p b_1 DomFil_{t-i} + \sum_{i=1}^p g_1 AK_{t-i} + \sum_{i=1}^p f_1 RD_{t-i} + \sum_{i=1}^p q_1 RGDP_{t-i} + e_t$$



Benchmark Forecasts

- These domestic or “first” filings are a strong indicator of subsequent filings at the EPO and used in the second stage. The specification is similar.

$$SEPOFil_t = b_0 + \sum_{i=1}^p b_1 DomFil_{t-i} + \sum_{i=1}^p g_1 SEPOFil_{t-i} + \sum_{i=1}^p f_1 EPOFil_{t-i} + \sum_{i=1}^p q_1 ERGDP_{t-i} + u_t \sum_{i=1}^n (X_i - \bar{X})^2$$

- Subsequent (or secondary) filings at the EPO are a function of past domestic filings, previous filings at the EPO, the size of the EPO market, and economic activity in Europe.



Benchmark Forecasts

Knowledge Production and Domestic Filing

$$DomFil_t^{US} = f_{US}(DomFil_{t-i}^{US}, AK_{t-1}^{US}, RD_{t-i}^{US}, GDP_{t-i}^{US})$$

$$DomFil_t^{JP} = f_{JP}(DomFil_{t-i}^{JP}, AK_{t-1}^{JP}, RD_{t-i}^{JP}, GDP_{t-i}^{JP})$$

$$DomFil_t^{EU} = f_{EU}(DomFil_{t-i}^{EU}, AK_{t-1}^{EU}, RD_{t-i}^{EU}, GDP_{t-i}^{EU})$$

(Subsequent) Filings at the EPO

$$SEPOFil_t^{US} = f_{US}(SEPOFil_{t-i}^{US}, DomFil_{t-1}^{US}, EPOFil_{t-i}^{Tot}, GDP_{t-i}^{US})$$

$$SEPOFil_t^{JP} = f_{JP}(SEPOFil_{t-i}^{JP}, DomFil_{t-1}^{JP}, EPOFil_{t-i}^{Tot}, GDP_{t-i}^{JP})$$

$$SEPOFil_t^{EU} = f_{EU}(SEPOFil_{t-i}^{EU}, DomFil_{t-1}^{EU}, EPOFil_{t-i}^{Tot}, GDP_{t-i}^{EU})$$

$$EPOFil_t^{OT} = f_{OT}(SEPOFil_{t-i}^{OT}, EPOFil_{t-i}^{Tot}, GDP_{t-i}^{EU})$$

$$EPOFil_t^{Tot} = EPOFil_t^{EU} + EPOFil_t^{JP} + EPOFil_t^{US} + EPOFil_t^{OT}$$



Benchmark Forecasts

The Domestic Filing Model - US

Specific model of LFDOM_US, 1968 - 2003

	Coeff	StdError	t-value
Constant	22.30339	5.68230	3.925
LFDOM_US_1	2.50788	0.58479	4.289
LFDOM_US_2	0.41380	0.17107	2.419
LFDOM_US_3	0.40580	0.15110	2.686
LRD3_US_1	0.80166	0.24127	3.323
LRD3_US_3	-1.29514	0.34318	-3.774
LAKD_US_1	-17.30588	5.35807	-3.230
LAKD_US_2	13.08304	4.50410	2.905
LAKD_US_4	0.71821	0.41603	1.726
dp9596	0.12257	0.02417	5.072
Trend	0.05670	0.01446	3.921



Benchmark Forecasts

The Domestic Filing Model - US

RSS	0.02714	sigma	0.03295
LogLik	129.42304	AIC	-6.57906
R^2	0.99424	Radj^2	0.99193
HQ	-6.41018	SC	-6.09520
T	36	p	11
		value	prob
Chow(1986:1)	3.1522	0.0515	
Chow(2000:1)	0.0387	0.9896	
AR 1-4 test	1.4191	0.2623	
ARCH 1-4 test	0.0876	0.9851	
hetero test	20.9696	0.3989	



Benchmark Forecasts

The Domestic Filing Model - US

Dynamic analysis - LongRun Estimates

LAKD_US		1.5058 >>> Greater than Unity
	SE	0.0712
LRD3_US		0.2120 >>> Elasticity .2
	SE	0.0573
dp9596		-0.0527
	SE	0.0196
Constant		-9.5826
	SE	0.9904
Trend		-0.0244
	SE	0.0039



Benchmark Forecasts

The EPO Total Model

Specific model of LF_TOT, 1982 - 2001

	Coeff	StdError	t-value	t-prob
Constant	7.83067	3.16504	2.474	0.0268
LF_TOT_2	0.78506	0.05666	13.855	0.0000
LAKT_US	3.05625	0.31196	9.797	0.0000
lrd31_eu_1	0.62121	0.08990	6.910	0.0000
lrd31_eu_2	-0.38245	0.07114	-5.376	0.0001
LGDP3_EU_1	-5.88867	0.98197	-5.997	0.0000
RSS	0.00796	sigma	0.02384	
LogLik	78.29397	AIC	-7.22940	
R^2	0.99822	Radj^2	0.99759	
HQ	-7.17108	SC	-6.93068	
T	20	p	6	
		value	prob	
Chow(2000:1)	0.2973	0.5948		
normality test	0.1066	0.9481		



Benchmark Forecasts

The EPO Total Model

Dynamic analysis

		LongRun
LAKT_US		14.2193
	SE	4.5099
lrd31_eu		1.1108
	SE	0.6108
LGDP3_EU		-27.3972
	SE	10.5798
Constant		36.4324
	SE	22.6889



Benchmark Forecasts

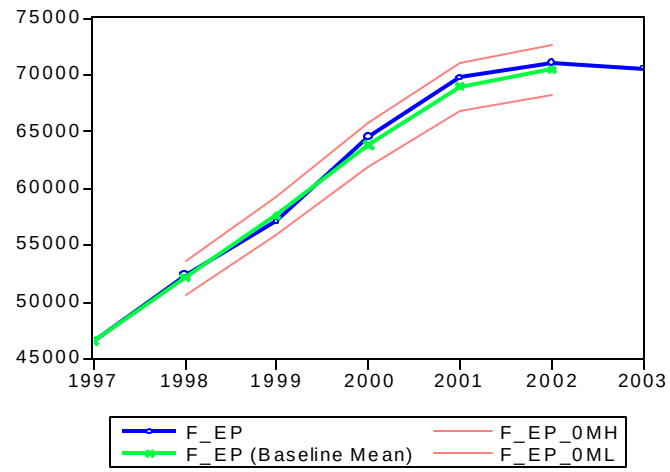
The Forecasts

- The aggregate and domestic models were solved dynamically in a stochastic simulation.
- 1000 repetitions Gauss-Seidel Method
- Actuals and Forecasts with Confidence Intervals
- Percent Deviations from Actual



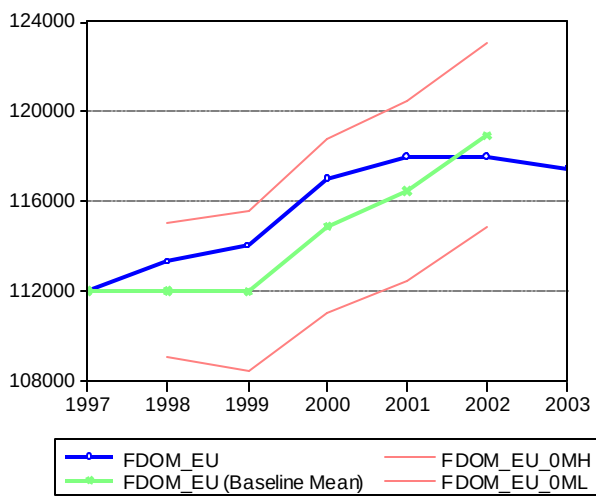
Benchmark Forecasts

Actual and Forecasts from the Aggregate Total Model



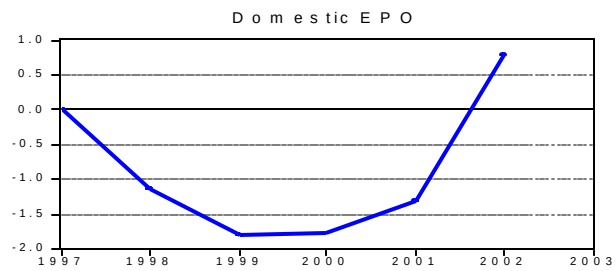
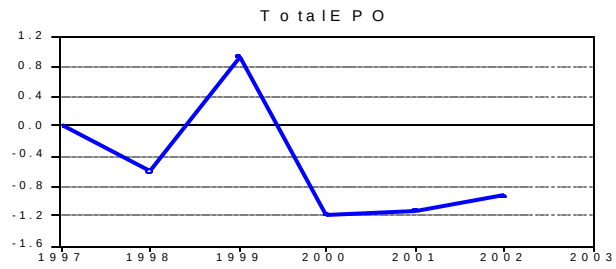
Benchmark Forecasts

Actual and Forecasts from the Aggregate Domestic Model



Benchmark Forecasts

Percent Deviations from Aggregate Model



Benchmark Forecasts