

March, 2012

MOODY'S ANALYTICS

# Improving Accuracy of Subnational Forecasting and Analysis through Historical Data Estimation

---

Prepared by

Megan McGee  
Assistant Director  
+610.235.5000

Karl Zandi  
Managing Director  
+610.235.5000

Steven G. Cochrane  
Managing Director  
+610.235.5000

---

# Improving Accuracy of Subnational Forecasting and Analysis through Historical Data Estimation

BY MEGAN MCGEE, KARL ZANDI AND STEVEN G. COCHRANE

**D**ecision-makers in industries such as real estate, utilities, retail trade, and consumer banking whose businesses are sensitive to local economic conditions have a fundamental need to understand subnational economies. But obtaining high quality historical data for subnational areas, the raw material of regional forecasting and analysis, presents challenges that are far greater than obtaining national-level data. Overcoming these obstacles is vital for timely identification of economic turning points, accurate forecasting, and robust model-building that businesses rely on for strategic planning and risk management.

Subnational data from public sources, if available, are difficult to convert to a usable form because of reporting lags, annual periodicity, definitional differences across nations, changing geographic boundaries, and industrial classification revisions. Attempts by public source aggregators such as Eurostat to create homogenized subnational datasets have failed to address the fundamental problems of highly lagged, low-frequency data with major series breaks. However, data estimation techniques overcome these challenges to create long, clean, quarterly historical time series that are extended years closer to the present day to represent current trends.

Through more than 20 years of experience specializing in subnational data, forecasting and analysis, Moody's Analytics has developed the techniques and infrastructure to deliver the most current and highest-frequency historical estimates of economic indicators for provinces, states, metropolitan areas and administrative districts worldwide. These estimates provide the most current information on subnational economies globally and are the foundation of our regional forecasts and analysis.

## Case Study: German Employment

Germany provides an excellent case study of how Moody's Analytics adds value

to subnational historical data. Employment data in Germany present the full spectrum of obstacles to creating a consistent national and subnational dataset. These obstacles include:

- » Multiple data sources reporting differing definitions of national employment
- » Inconsistency between the national and subnational data
- » Frequent changes in subnational geographic definitions with no officially provided revisions to prior data
- » Highly lagged historical data
- » Varying update frequencies and periodicities of the data

Moody's Analytics begins creating a consistent national and subnational dataset by researching all available national data sources and surveys. For Germany, this revealed multiple sources reporting data using either a "resident" or "domestic" concept of employment. The resident concept is employment by place of residence, while the domestic concept is employment by place of work. Nonresidents working in Germany and residents of Germany working abroad account for the difference between these concepts.

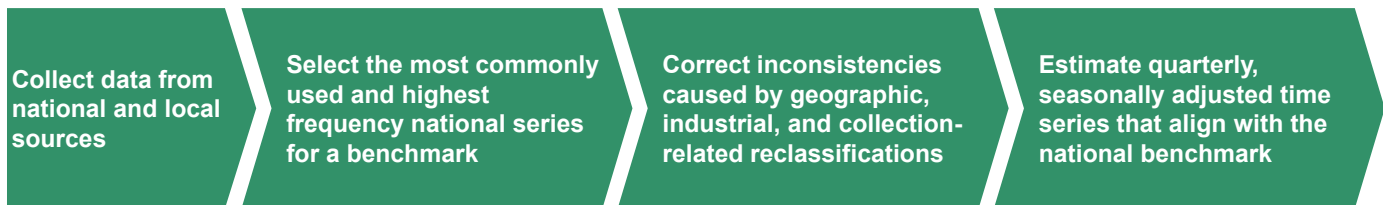
Within these two broad concepts, reported employment figures differ among Eurostat, the Federal Statistical Office of

Germany (FSO); the Federal Employment Agency of Germany (FEA); and Deutsche Bundesbank, Statistics Department (DBB), among other sources. Moody's Analytics selected the broadest definition of employment used by these sources—national accounts employment. The employment accounts integrate approximately 50 statistics obtained through different reporting channels. These statistics include employees subject to social insurance contributions and marginal part-time workers as compiled by the FEA; public service personnel statistics; results of the micro census; and other information from various institutions, such as monthly reports from the Federal Ministry of Defense on the number of soldiers.

The employment accounts are released monthly, quarterly, and as annual averages by the FSO. The high-frequency monthly figure is released about 30 days after the end of the reference period by the FSO, and DBB seasonally adjusts this series prior to publishing. Moody's Analytics uses the monthly figures released by DBB as the benchmark series for the German employment estimates.

The monthly figures meet the need for high-frequency, up-to-date information, but lack the breadth of the quarterly and annual

**Flow Chart: Sub-national Historical Data Estimation Process**



figures. Therefore, the monthly figures are adjusted by FSO several times to coincide with the more accurate quarterly and annual figures. Additionally, benchmark revisions are performed as part of the national accounts every several years, the most recent in April 2005.

**Consistent Subnational Geography**

After selecting the highest frequency and broadest concept available for the benchmark series nationally, Moody's Analytics begins researching the geographies and data availability subnationally. For Germany, this research revealed regular geographic redefinitions and low-frequency, highly-lagged subnational datasets.

Under the Nomenclature of Territorial Units for Statistics (NUTS), districts in the states of Germany known as Kreise, or NUTS 3, are regularly revised. Most sources, including Eurostat, rarely revise historical data after these geographic redefinitions. Therefore, most datasets contain time series breaks for the affected geographies. To compensate for the breaks, Moody's Analytics adjusts the historical series. If a newly created area is composed of several former areas, these areas are aggregated. If the newly created areas are created by splitting areas, Moody's Analytics either applies a known ratio (e.g. population) over the history of the former area or uses a consistent ratio calculated after the break.

Chart 1 illustrates an example of a geographic redefinition in Germany in two areas that are contained in the Berlin metropolitan area. To take one example, Potsdam was revised in 2003 to include part of Potsdam-Mittelmark. Eurostat has not yet revised these areas' history and may or may not with the release of new NUTS classifica-

tions. Moody's Analytics revises the break in the series using a constant rate for 2004 applied to the history of these series.

**Time Series Estimation**

For all German subnational areas, estimation begins with the annual national accounts employment figures released by the Federal Statistical Office. The subnational employment accounts are consistent with the monthly employment figures released by Deutsche Bundesbank, Statistics Department, for Germany when averaged.

Moody's Analytics then utilizes subnational high-frequency data from the Federal Employment Agency to produce high-frequency estimates subnationally. The FEA reports two surveys that the FSO uses as its main inputs to employment accounts: Employees Contributing to Social Security (Sozialversicherungspflichtig Beschäftigte) and Marginally Employed (Geringfügig Beschäftigte). The marginally employed are workers whose salaries are consistently less than €400 a month; therefore, they do not contribute to social security. When aggregated these two surveys exclude freelance employees, workers assisting family members, soldiers with enlistment contracts of more than two years, conscripted members of the military with no prior employment contributing to social security, and tenured public employees such as teachers and judges. However, these two surveys aggregated repre-

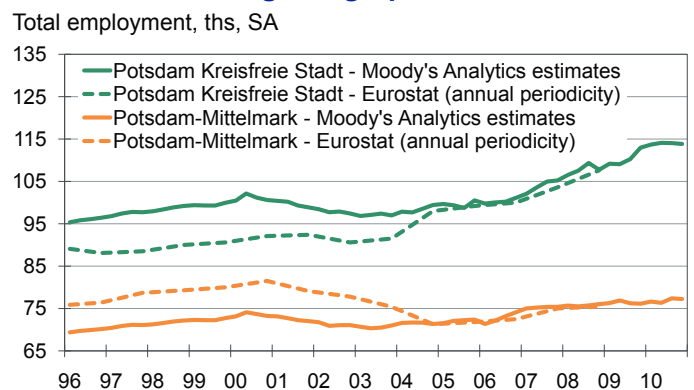
sent approximately 80% of employment in Germany and are therefore a good proxy for overall employment.

The FEA surveys are reported quarterly, without seasonal adjustment. Moody's Analytics seasonally adjusts them and applies their quarterly pattern to the annual employment accounts subnationally to produce higher frequency estimates. Moody's Analytics then uses their quarterly growth rates to extend the series through the latest quarter of FEA data.

Moody's Analytics creates consistency between the national and subnational employment estimates by squeezing subnational areas using a top-down approach. This ensures that subnational geographies aggregate correctly to national. The top-down squeeze is applied first to the states to aggregate to the national level, then the regions to aggregate to the states, and finally to the districts to aggregate to the regions.

As a final step, Moody's Analytics produces employment estimates for metropolitan areas by aggregating the NUTS 3 districts to metropolitan areas as specified by the Urban Audit from Eurostat. Moody's Analytics cre-

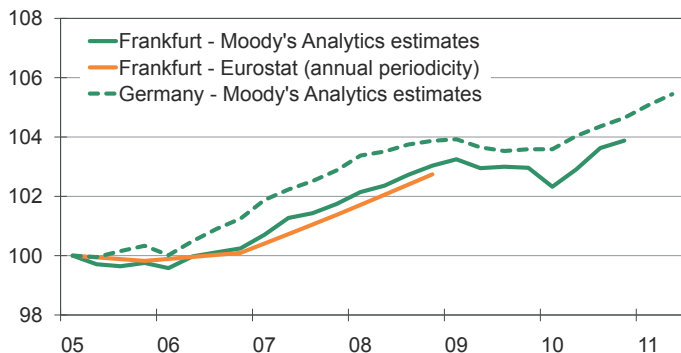
**Chart 1: Correcting Geographic Series Breaks**



Sources: Eurostat, Deutsche Bundesbank, FSO, FEA, Moody's Analytics estimates

**Chart 2: Capturing Turning Points in Frankfurt**

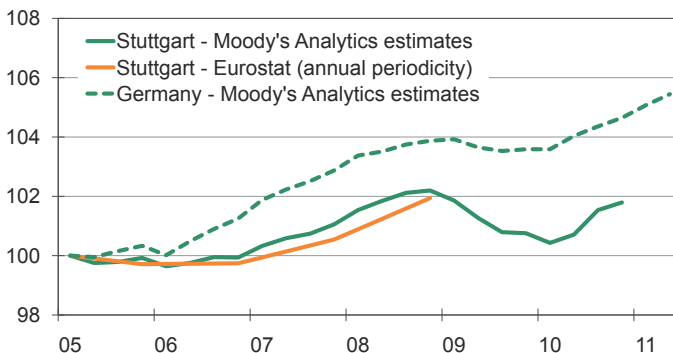
Total employment, 2005Q1 = 100, SA



Sources: Eurostat, Deutsche Bundesbank, FSO, FEA, Moody's Analytics estimates

**Chart 3: Capturing Turning Points in Stuttgart**

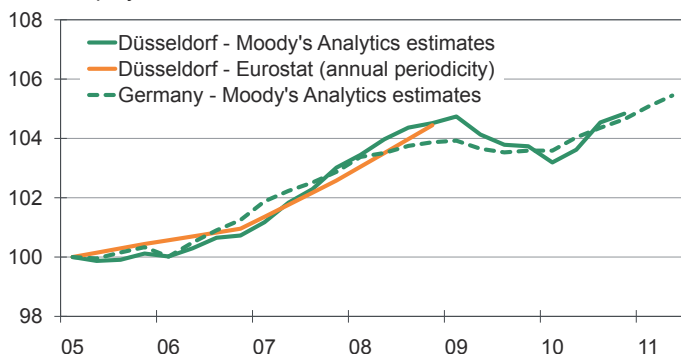
Total employment, 2005Q1 = 100, SA



Sources: Eurostat, Deutsche Bundesbank, FSO, FEA, Moody's Analytics estimates

**Chart 4: Capturing Turning Points in Düsseldorf**

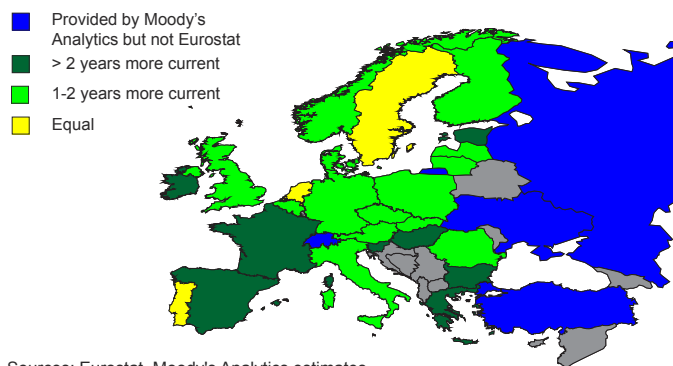
Total employment, 2005Q1 = 100, SA



Sources: Eurostat, Deutsche Bundesbank, FSO, FEA, Moody's Analytics estimates

**Chart 5: Data Are Several Years More Current**

Total employment data, Moody's Analytics vs. Eurostat



Sources: Eurostat, Moody's Analytics estimates

ates metropolitan area aggregates globally to analyze and forecast cohesive economic units and facilitate comparisons of subnational areas across countries and continents.

**Global Historical Data Challenges**

The challenges faced to produce a high quality subnational historical dataset for Germany are not unique. Moody's Analytics addresses a similar set of challenges with U.S. state and metropolitan areas, synthesizing three separate employment surveys in much the same way to overcome the individual shortcomings of each dataset.

We address these challenges across nearly all of Europe as well. We produce quarterly historical estimates for more than 130 metropolitan areas in Europe, where more than half the countries provide only annual subnational data and almost all contain series breaks because of geographic or industrial classification redefinitions without

full revisions of prior data releases from the original sources.

**Improved Analysis and Forecasting**

Cleaning, extending, and increasing the frequency of historical subnational time series benefits those analyzing and modeling local economies in several ways, including:

- » Identifying turning points in local economies sooner with real data
- » Improving forecast accuracy
- » Enhancing economic models with longer time series and quarterly periodicity for economic drivers

In the German employment example, our estimation techniques extend the length of historical time series for NUTS 3 total employment forward by two whole years.<sup>1</sup>

<sup>1</sup> The lag in Eurostat data versus Moody's Analytics estimates was measured on January 20, 2012. The lag may vary over time depending on the relative timing between Eurostat and national source reporting.

Because of this lag, the Eurostat data miss several turning points in the economies of several key German metropolitan areas such as Frankfurt (see Chart 2), Stuttgart (see Chart 3), and Düsseldorf (see Chart 4), which the Moody's Analytics estimates capture.

The ability of Moody's Analytics historical estimates to capture subnational economic turning points years before Eurostat can be expected not just in Germany, but across Europe<sup>2</sup> (see Chart 5). Where coverage overlaps, Moody's Analytics' data are typically one to three years more current than Eurostat's.

The quality of historical data also has profound implications for forecast accuracy. From our experience in forecasting subnational economies, improving the accuracy of historical data can reduce forecast error by

<sup>2</sup> Lag in Eurostat data versus Moody's Analytics estimates was measured on January 20, 2012.

50% or more.<sup>3</sup> The combined effects of data estimation to extend historical time series, real-time historical data updates, and month-

---

<sup>3</sup> Based on analysis of the effect of benchmark revisions on forecast accuracy for the total employment of U.S. states and metropolitan areas. See *Regional Financial Review* Regional Forecast Accuracy articles for years 1994-2002.

ly forecast updates ensure that Moody's Analytics subnational forecasts are always based on the most current data possible, reducing forecast error due to historical data.

Finally, through historical estimation, we are also able to build more robust economic models for metropolitan economies

from a larger and more detailed dataset. The historical time series from which we derive our forecast models are longer going both forward and backward in time to cover more business cycles than otherwise possible, with richer information content provided through quarterly periodicity.

# About the Authors

## Megan McGee

Megan McGee is assistant director of Data Services for Moody's Analytics. She is responsible for developing and maintaining the company's estimated data products, including global national and sub-national historical estimates and the U.S. detailed employment, output, and wages datasets. Megan holds a bachelor's degree from Pennsylvania State University.

## Karl Zandi

Karl Zandi is managing director of Data Services for Moody's Analytics. He is responsible for all commercial and technical aspects of the company's data operations, leading a team of more than 20 data specialists located in Prague, Sydney, and West Chester. Karl holds a bachelor's degree from West Chester University.

## Steven G. Cochrane

Steven G. Cochrane is managing director of Moody's Analytics. Steve oversees the U.S. regional forecasting service and directs the research and development activities of the research staff. He oversees the forecasts for all 50 states and developed the Global Cities service. He also edits the Regional Financial Review, the monthly publication that analyzes U.S. macro, regional, industry and international trends. An analyst with Moody's Analytics since 1993, Steve has been featured on Wall Street Radio, the PBS NewsHour, and CNBC. He earned a PhD in regional science at the University of Pennsylvania, a master's degree at the University of Colorado at Denver, and a bachelor's degree at the University of California at Davis.

## More on Global Economies from Moody's Analytics

### GLOBAL METRO AREA FORECASTS

Insights into the current and expected economic conditions of metropolitan areas worldwide.

[www.economy.com/globalmetros](http://www.economy.com/globalmetros)

### GLOBAL DATABASE

More than 250 million national and subnational historical time series, covering more than 180 countries.

[www.economy.com/globaldata](http://www.economy.com/globaldata)

U.S./CANADA  
+1 866.275.3266

EMEA  
+44 (0) 20.7772.1646

ASIA/PACIFIC  
+61 2 9270 8111

OTHER LOCATIONS  
+1 610.235.5299

Email: [help@economy.com](mailto:help@economy.com)

# About Moody's Analytics

## Economic & Consumer Credit Analytics

Moody's Analytics helps capital markets and credit risk management professionals worldwide respond to an evolving marketplace with confidence. Through its team of economists, Moody's Analytics is a leading independent provider of data, analysis, modeling and forecasts on national and regional economies, financial markets, and credit risk.

Moody's Analytics tracks and analyzes trends in consumer credit and spending, output and income, mortgage activity, population, central bank behavior, and prices. Our customized models, concise and timely reports, and one of the largest assembled financial, economic and demographic databases support firms and policymakers in strategic planning, product and sales forecasting, credit risk and sensitivity management, and investment research. Our customers include multinational corporations, governments at all levels, central banks and financial regulators, retailers, mutual funds, financial institutions, utilities, residential and commercial real estate firms, insurance companies, and professional investors.

Our web and print periodicals and special publications cover every U.S. state and metropolitan area; countries throughout Europe, Asia and the Americas; and the world's major cities, plus the U.S. housing market and other industries. From our offices in the U.S., the United Kingdom, and Australia, we provide up-to-the-minute reporting and analysis on the world's major economies.

Moody's Analytics added Economy.com to its portfolio in 2005. Its economics and consumer credit analytics arm is based in West Chester PA, a suburb of Philadelphia, with offices in London and Sydney. More information is available at [www.economy.com](http://www.economy.com).

---

© 2012, Moody's Analytics, Inc. and/or its licensors and affiliates (together, "Moody's"). All rights reserved. ALL INFORMATION CONTAINED HEREIN IS PROTECTED BY COPYRIGHT LAW AND NONE OF SUCH INFORMATION MAY BE COPIED OR OTHERWISE REPRODUCED, REPACKAGED, FURTHER TRANSMITTED, TRANSFERRED, DISSEMINATED, REDISTRIBUTED OR RESOLD, OR STORED FOR SUBSEQUENT USE FOR ANY PURPOSE, IN WHOLE OR IN PART, IN ANY FORM OR MANNER OR BY ANY MEANS WHATSOEVER, BY ANY PERSON WITHOUT MOODY'S PRIOR WRITTEN CONSENT. All information contained herein is obtained by Moody's from sources believed by it to be accurate and reliable. Because of the possibility of human and mechanical error as well as other factors, however, all information contained herein is provided "AS IS" without warranty of any kind. Under no circumstances shall Moody's have any liability to any person or entity for (a) any loss or damage in whole or in part caused by, resulting from, or relating to, any error (negligent or otherwise) or other circumstance or contingency within or outside the control of Moody's or any of its directors, officers, employees or agents in connection with the procurement, collection, compilation, analysis, interpretation, communication, publication or delivery of any such information, or (b) any direct, indirect, special, consequential, compensatory or incidental damages whatsoever (including without limitation, lost profits), even if Moody's is advised in advance of the possibility of such damages, resulting from the use of or inability to use, any such information. The financial reporting, analysis, projections, observations, and other information contained herein are, and must be construed solely as, statements of opinion and not statements of fact or recommendations to purchase, sell, or hold any securities. NO WARRANTY, EXPRESS OR IMPLIED, AS TO THE ACCURACY, TIMELINESS, COMPLETENESS, MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OF ANY SUCH OPINION OR INFORMATION IS GIVEN OR MADE BY MOODY'S IN ANY FORM OR MANNER WHATSOEVER. Each opinion must be weighed solely as one factor in any investment decision made by or on behalf of any user of the information contained herein, and each such user must accordingly make its own study and evaluation prior to investing.

**CONTACT US**

For further information contact us at a location below:

**U.S./CANADA**  
+1.866.275.3266

**EMEA**  
+44.20.7772.5454

**ASIA/PACIFIC**  
+852.3551.3077

**OTHER LOCATIONS**  
+1.610.235.5299

Email us: [help@economy.com](mailto:help@economy.com)  
Or visit us: [www.economy.com](http://www.economy.com)

Copyright © 2012, Moody's Analytics, Inc. All Rights Reserved.