



German Council for Social
and Economic Data (RatSWD)

www.ratswd.de

RatSWD

Working Paper Series

Working Paper

No. 64

Improvements and Future Challenges for the Research Infrastructure in Public Finance

Thiess Buettner

March 2009

Working Paper Series of the Council for Social and Economic Data (RatSWD)

The *RatSWD Working Papers* series was launched at the end of 2007. Since 2009, the series has been publishing exclusively conceptual and historical works dealing with the organization of the German statistical infrastructure and research infrastructure in the social, behavioral, and economic sciences. Papers that have appeared in the series deal primarily with the organization of Germany's official statistical system, government agency research, and academic research infrastructure, as well as directly with the work of the RatSWD. Papers addressing the aforementioned topics in other countries as well as supranational aspects are particularly welcome.

RatSWD Working Papers are non-exclusive, which means that there is nothing to prevent you from publishing your work in another venue as well: all papers can and should also appear in professionally, institutionally, and locally specialized journals. The *RatSWD Working Papers* are not available in bookstores but can be ordered online through the RatSWD.

In order to make the series more accessible to readers not fluent in German, the English section of the *RatSWD Working Papers* website presents only those papers published in English, while the German section lists the complete contents of all issues in the series in chronological order.

Starting in 2009, some of the empirical research papers that originally appeared in the *RatSWD Working Papers* series will be published in the series *RatSWD Research Notes*.

The views expressed in the *RatSWD Working Papers* are exclusively the opinions of their authors and not those of the RatSWD.

The RatSWD Working Paper Series is edited by:

Chair of the RatSWD (2007/ 2008 Heike Solga; 2009 Gert G. Wagner)

Managing Director of the RatSWD (Denis Huschka)

Improvements and Future Challenges for the Research Infrastructure in Public Finance

Thiess Buettner

Abstract

The paper briefly surveys available data sources and discusses future developments relevant for empirical public finance in Germany. It starts from the notion that public finance deals with decisions of various agents, not only governments, but also private households and firms. Therefore, empirical research needs different types of data. Government decisions are to some extent captured in terms of the budgetary statistics, even though these statistics have shortcomings with regard to the quality of public service provisions and the revenue instruments. To study the decisions of the other agents individual level data is required. While some progress has been made, recently, the combination of various datasets at the individual level is a key priority.

Keywords: Empirical Research; Public Finance; Budgetary Statistics; Revenue Statistics; Micro-level Data; Taxpayer Data

1. Research Questions

Public finance is concerned with decisions of collective agents and institutions and with the impact of those decisions on the economy and individual agents. This implies that depending on the specific research topic rather different types of data are used and may have to be combined.

Empirical research in public finance traditionally addresses the decision making of the public sector itself. This includes substantial efforts to monitor the developments in the public sector in terms of the budget as well as with regard to service provision and inputs. In most countries the public sector shows a marked vertical and horizontal structure such that research addresses all sorts of governments including national and federal governments, state governments, as well as local governments and separate bodies such as school districts and public enterprises.

With regard to policy areas, public finance tends to take a comprehensive view. Research in areas where policy implementation takes specific forms and includes private institutions has developed into own subfields. This includes areas such as Health and Education. While some of the research in these areas has close connections to public finance in general, this paper will not discuss the empirical research infrastructure in these areas.

A significant part of empirical research in public finance focuses on the impact of instruments of public policies on the economy – such as taxes and various types of government spending including subsidies and transfers. A particular focus is on the impact of those instruments on individual agents, such as households, and on firms. This usually requires data about individual agents. At the same time, however, to study long-term implications of policies not only cross-sectional but also longitudinal data are required.

Due to the complexity, research on taxation is often concerned with simulating rather than testing for tax effects. The corresponding simulation models need detailed information about the various components of the tax accounts and, hence, are ideally based on micro-level data for taxpayers.

Given the large share of resources expended by the public sector, empirical research is also concerned with the macroeconomic consequences of fiscal policies at an aggregate level such as the national, regional, or state levels. There is also empirical research that aims to give a comprehensive picture of the economic consequences of government policies using general-equilibrium simulation models. These models require detailed information not only about the public sector but also about household and firm sectors and may also utilize input-output tables. Moreover, these models usually employ various parameters that originate in previous empirical research.

Empirical research in public finance, however, is not only concerned with different types of agents, it also faces specific measurement problems with regard to government decisions and policies. Often, research is concerned with data on government spending or revenues. However, in many circumstances empirical research benefits from using more detailed information on specific government policies. This holds in particular in the context of taxation where actual policies are concerned with determining very specific parameters such as statutory tax rates, tax brackets, or tax incentives. Generally, detailed knowledge of the law and its implementation is required. Measurement issues are also important with regard to the supply side of the public sector where the use of expenditure data is often not sufficient to capture government policies, if the analysis is concerned with public service provision. As quantity and quality of public service provision are difficult to capture, research often takes resort to using survey data where respondents assess the supply of public services.

2. Status Quo: Data Bases and Access

Given the various types of research questions a useful way to structure a discussion of databases might distinguish the agents whose decisions are under consideration as well as the type of policy under consideration such as taxes, public service provision, or social policies.

2.1 Data on Governments

The basic datasource for empirical research is the fiscal accounts that capture expenditures, revenues as well as information on stocks such as government debt or assets. With regard to Germany the statistical offices of the federation and the states offer a broad set of detailed data covering public spending detailed by types of expenditure and categories of revenue. These statistics refer to the various fiscal tiers (federal, state, county, and municipality) and

some important *parafisci* such as social insurance. Quarterly data are provided about three months after the end of the respective quarter. Detailed data for functions of government are based on the annual accounts that are available about two years after the respective year. Budget information is augmented by aggregate tax revenue statistics that report revenue detailed by specific taxes. Furthermore the standard set of statistics includes information about the stocks of debt. The monitoring of government activities by the statistical offices further includes information about employment in the public sector. Finally, information is provided about the annual accounts of state owned enterprises.

Data access is easy for federal and state-level data as well as for the consolidated budget of the public sector: in all cases information is available on the website of the federal statistical office. However, with regard to the state level not all statistics are provided as part of the standard program: for instance, data detailed by both type of expenditure and function of government need separate requests. While a virtue of the German system of fiscal federalism is that the fiscal classification used for counties and municipalities is almost identical across states, detailed data on government below state level are only available for individual states subject to the approval of the respective state's statistical office.¹

Given the large expenditure share of welfare, several statistics of the statistical offices focus on various programs, and also the Federal Ministry of Labour and Social Affairs provides comprehensive statistics (*Sozialbudget*).

For several major taxes, the statistical offices hold individual tax files that are used to provide some more detailed information and are made available for research as scientific use files. Combination into panel data is possible if based on tax identifiers. However, with most taxes the data are triennial starting with 1992 or 1995. Data access is restricted to the research centers of the statistical offices and may be further restricted by remote processing of routines.

While the fiscal variables refer to the executive branch of the government, information about the legislature is also provided by the statistical offices. At the website <http://www.bundeswahlleiter.de> detailed data on the results of federal, state, and local elections can be downloaded. Information about annual and medium term budget planning is provided through the federal department of finance (BMF) or through the finance departments

1 An exception is a study by Borck, Caliendo, and Steiner (2007) that uses a comprehensive dataset for all German municipalities.

of the states. Data on auditing is available for the federal and the state level, only. Results of the auditing of lower level governments are generally not available.

With regard to quantity and quality of public services provided by the various governments the data supplied by the statistical office is rather limited. For some functions of government specific statistics exist. For instance, statistics on higher education provide data on enrollment at universities and the “universities of applied sciences” (Fachhochschulen) by university and field of study including further information about the background of students. However student test scores with regard to schooling as assembled in the OECD’s PISA initiative are not provided at state level in a way that allows meaningful cross-state comparisons. Since education is the key responsibility of the state governments in Germany this restriction is a severe limitation for empirical research in Germany.

At the municipal level the statistical yearbook for the cities (edited by the German Association of Cities) provides some further information on the supply of government services, however, this data focuses on larger cities. Research that is concerned with the supply of public services at the local level might take resort to survey data, where, however, the number of respondents is often small. An exception is the ”Perspektive Deutschland” where waves four (2004/2005) and five (2005/2006) provide survey responses for several aspects of local living conditions including public services at county level. Data access is provided through the Zentralarchiv (ZA) in Cologne.

2.2 Household Data

An important part of empirical research in public finance is concerned with the impact of public policies on household decisions such as consumption, labor supply, or location. For this purpose, all sorts of household data are used such as SOEP, Mikrozensus, or EVS. The latter is particularly interesting as it offers some direct information about taxes paid whereas the SOEP employs imputed values (Becker et al., 2002).

Given the importance of specific institutional details of the tax code, research often uses taxpayer panels (see below) even if these have limited information about household characteristics. This includes the IAW-Einkommensteuerpanel (Gottfried und Schellhorn, 2001) that builds on individual tax information in the state of Baden-Wuerttemberg. More recently, the federal statistical office provides an annual taxpayer panel, a project that

provides taxpayer data on an annual basis starting in 2001 (see Kriete-Dodds and Vorgrimler, 2007). While the taxpayer panel can only be used via controlled remote processing, the triennial micro-level income tax statistics is available as scientific use file (FAST) within the research centers of the statistical offices (FDZ).²

Another important area of research is concerned with the consequences of social policy on individual choice. However, key issues in this context such as distribution or labor market participation suggest to discuss data availability and the conditions for empirical analysis in the context of poverty (see Hauser, 2009) and labor market research (see Bender and Möller, 2008, Schneider, 2008).

2.3 *Firm Data*

To study the impact of government policies on firm decisions a large body of research utilizes firm-level data that capture investment, financial structure, and many other dimensions of firm decisions. However, financial statement data as provided by Hoppenstedt or Creditreform (DAFNE) usually report tax payments that capture not only the tax burden or tax incentives but are also reflecting firm performance and/or tax planning. The resulting problem of the endogeneity of tax variables has made it difficult to identify the role of the tax system for investment or the financial structure of firms. Rather than using tax payments, research might exploit differences between firms that lead to differences in taxation due to the specifics of the tax law, perhaps related to legal form or firm size.

In the last decade, however, empirical research has been more successful in addressing these issues by employing data for multinational firms operating in different countries. The advantage here is that policies including tax policies show marked variation across countries that can be exploited for identification purposes. As a consequence, much research has been concerned with multinational data. Financial statement data for German and European multinationals or multinationals operating in Europe are provided by commercial providers such as Bureau von Dijk's Amadeus database. A unique datasource for studying multinationals is the Bundesbank's Midi database that currently provides annual firm-level panel data for the period 1996 to 2004. The collection of the data is prescribed by German law, which determines reporting mandates for international transactions (Lipponer, 2006). A

² Here and in the following FDZ refers to the research centers of the Federal statistical office and the statistical offices of the States (Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder).

shortcoming of the Midi dataset is that it provides limited information about the parent companies of domestic affiliates and German parents of foreign affiliates.

Alternatively, research is concerned with data that exploits institutional variations across regions. In the context of company taxation, for instance, many studies exploit the local variation in the local business tax (Gewerbesteuer). Research opportunities are provided by the corporate balance sheet database (Jahresabschlussdaten) of the Bundesbank (Stoess, 2001), available within the bank's research center, by the IAB establishment panel, where scientific use files are provided through the research institute of the federal employment agency, and by the taxpayer data for the local business tax. The latter is a triennial micro-level dataset that currently provides information for 1998 and 2001. It can be accessed within the research centers of the statistical office, though data access is further restricted by controlled remote processing.

Given the difficulties in identifying tax effects, research is also concerned with setting up simulation models. As with the analysis of household decisions micro-level tax statistics are particularly helpful for this purpose. In addition to the local business tax, the triennial micro-level tax statistics for the corporation tax, and, with regard to the unincorporated firms, also the personal income tax statistic can be used – all of them are provided within the research centers of the statistical offices (FDZ).

2.4 Tax Policy & Institutional Data

Typical for a large part of empirical research in public finance is a detailed characterization of institutions including tax systems, social security systems, or specific laws that govern government policies. At the international level, however, some supra-national bodies such as the OECD and the European Commission exist that provide data on tax systems and institutional characteristics of countries such as the vertical structure of the public sector.

Generally speaking data collection is easier for sub-national entities. For instance, information about the tax burden associated with the local business tax (Gewerbesteuer) and the land tax (Grundsteuer) is provided by the statistical offices, at least at the level of counties.³ However, the effective tax burden on land is not known due to the substantial

³ Common practice is to report weighted averages of municipal tax rates where the weights correspond to the municipal government's tax revenue. However, this practice is problematic. To see this, consider the weighted collection rate (Hebesatz) for a set of municipalities:

discrepancy between market value of land and the assessed value. While land is assessed for the estate and gift tax (Erbschaftsteuer) according to market prices since the mid of the nineties, information about the assessment is not provided by the statistical offices.

3. Future Developments

An important issue for future development is the combination of various datasources. This refers in particular to corporate and personal income taxation where taxpayer panel data so far have not been merged between taxes. However, even at the level of the corporation the tax burden consists of local business taxes as well as of corporate taxes. Moreover, a combination of taxpayer data with other firm level data could yield substantially improved datasets where firm decisions as well as firm-specific conditions could be modelled much more precisely. The recent KombiFiD initiative of the federal statistical office, the research institute of the federal employment agency (IAB), and of the Deutsche Bundesbank aiming at providing those combined datasets is to be strongly appreciated (see Bender et al.,2007).

In particular, the combination with the Bundesbank MiDi data would vastly improve conditions for empirical research. Another promising data combination project is the EBDC initiative of Munich University together with the ifo Institute that aims at combining ifo-firm survey data (Becker and Wohlrabe, 2008) with commercial financial statement data including Amadeus and Hoppenstedt. An interesting aspect of this project is that it relies on randomized record linkage.

A data combination that would help to address important issues with regard to the tax system is a combination of individual and firm-level taxpayer data. This would be an important step towards creating a reliable and comprehensive empirical basis for research on tax policy and reforms. The federal department of finance (BMF) has started an initiative in this direction. In this context, it should be noted that the growing complexity of tax issues has led other countries to set up micro-simulation models that are used for revenue estimation and also revenue forecasting purposes. However, currently no such attempts have materialized in

$$\frac{\sum_{i=1}^N hiGi}{\sum_{i=1}^N Gi}$$

where G_i is the revenue at a standardized collection rate (Grundbetrag). Since G_i is a declining function of the own collection rate and an increasing function of the tax rates of other municipalities, municipalities with high tax rates tend to receive a smaller weight. As a consequence, using the weighted average tends to yield biased results: tax increases are underestimated, tax decreases are overestimated.

Germany.

With regard to the analysis of the finances of subnational jurisdictions the traditional financial accounting system is subject to change. As of 2009 Northrhine-Westfalia, the largest German state, introduces a new system of accounts “Neues Kommunales Finanzmanagement” (NKF) that replaces the current cash-based accounting. While this might systematically improve the information about controlled enterprises and liabilities some new problems regarding assessment and valuation are coming up. Moreover, the data series will suffer from an important structural break.

As the current analysis of municipal finances has to take resort to the last 1987 census, another important development for empirical research at the local level is the new census that is planned for 2011. While this update is important, it should be noted that currently access to census data at municipal level is difficult as data access is restricted by the state statistical offices.

4. Future Developments: European and International Challenges

Much research in public finance is concerned with the consequences of international economic and political integration for public policies. This includes cross-border flows not only of goods and services but also migration, factor movements, capital flows, and the emergence of multinational enterprises. Even if those cross-border issues are of particular importance in the European context, there is only very limited information available. Research so far has centered around specific datasets, many of which are subject to important qualifications. To merge those datasets with other more standard datasets of households and firms would substantially improve the conditions for empirical research using German data. Therefore, initiatives like KombiFiD are very welcome.

5. Conclusions and Recommendations

Empirical research in public finance that aims at monitoring and assessing the budgetary performance has access to a rich body of financial accounts that enables to assess federal and state budgets. While the statistics are also available for lower level governments, data access for municipalities and related governments is unduly restrictive.

With regard to statistics on major taxes the statistical offices of the states together with the federal statistic office have recently improved conditions for empirical research by providing micro-level taxpayer data. This is a major achievement. Whether the restrictions in data access and the limited information about the background of tax payers in these statistics constitute a significant obstacle for the exploitation of this data remains to be seen.

Attempts to combine different data for research purposes are greatly appreciated. However, data combination should not be confined to provide firm-level data; it would also be important to combine different taxpayer statistics in order to get a comprehensive and consistent data source for the empirical analysis of the tax system.

With regard to the supply of public services, data availability for research is not satisfying. To some extent this reflects the fundamental problem of measuring public services – expenditure data offers only very limited information about the quantity and quality of public service provision. Some progress has been made in specific areas of government policies such as social policy. However, in other areas, such as public education, information about quality that is available, in principle, has not been made available for research.

References:

- Becker, I. , J. Frick, M. M. Grabka, R. Hauser, P. Krause, and G. G. Wagner, 2002, A Comparison of the Main Household Income Surveys for Germany: EVS and SOEP, in: R. Hauser und I. Becker, Reporting on Income Distribution and Poverty. Perspectives from a German and European Point of View, Heidelberg: Springer, 55–90.
- Becker, S. O., and K. Wohlrabe, 2008, Micro Data at the Ifo Institute for Economic Research - The “Ifo Business Survey” Usage and Access, *Schmollers Jahrbuch* 128(2), 307–319.
- Bellmann, L., 2002, Das IAB-Betriebspanel: Konzeption und Anwendungsbereiche, *Allgemeines statistisches Archiv* 86 (2), 177–188.
- Bender, S. and J. Möller, 2008, Developing the Research Data Infrastructure for the Social and Behavioral Sciences in Germany and Beyond Special topic: Labor Market I, Manuscript, Berlin.
- Bender, S., J. Wagner, and M. Zwick, 2007, KombiFiD - Kombinierte Firmendaten für Deutschland, FDZ-Arbeitspapier Nr. 21, Wiesbaden.
- Borck, R., M. Caliendo, and V. Steiner, 2007, Fiscal Competition and the Composition of Public Spending: Theory and Evidence, *FinanzArchiv* 63(2), 264–277.
- Gottfried, P. and H. Schellhorn, 2001, Das IAW-Einkommensteuerpanel und das Mikrosimulationsmodell SIMST, IAW-Diskussionspapier, Tübingen.
- Hauser, R., 2009, Household income, poverty and wealth, RATSWD Working Paper 53, Berlin.
- Kriete-Dodds, S., and D. Vorgrimler, 2007, Das Taxpayer-Panel der jährlichen Einkommensteuerstatistik, *Wirtschaft und Statistik*, 1 (2007), 77–85.
- Lipponer, A., 2006, Microdatabase direct investment - MiDi. A brief guide, Bundesbank Working Paper, Frankfurt.
- Schneider, H., 2008, Developing the Research Data Infrastructure for the Social and Behavioral Sciences in Germany and Beyond Special topic: Labor Market II, Manuscript, Berlin.
- Stoess, E., 2001, Deutsche Bundesbank’s Corporate Balance Sheet Statistics And Areas of Application, *Schmollers Jahrbuch*, 121, 131–137.