

**Review of “Measurement of Economic Variables”**  
**Dissertation submitted by Dana Hájková, November 2006**

Introduction

The “economics of measurement” has been a subject that has received too little attention in the economics discipline in recent decades. The incentives within the profession to work on measurement issues are not great these days. It is therefore almost a sign of great courage if someone undertakes such painstaking work on economic statistics, in particular as part of a dissertation. The candidate is therefore to be applauded for the excellent piece of work (or rather three pieces of work) on the topic of measurement of economic performance. This is all the more important because the topics addressed, namely measurement of variables on economic growth and sources of growth, are increasingly obscured due to the greater complexity of the economic growth process in particular in advanced countries. The increased importance of intangibles, whether it is service output, software inputs or human capital, puts huge challenges on the data, and requires serious research.

Ms. Hájková’s is clearly in the top of the class of work on “economics of measurement”, and is clearly worthy of a doctorate degree. All three essays represent a careful analysis of the basic measurement issues and great attention for detail but with an ability to keep an eye on “the big picture” and not get completely lost in the woods. The challenges to measurement also demand for a strong theoretical basis to measurement. While the thesis is strong on attention for empirics, it is at places weaker on questioning the theoretical basis of measurement. Are we measuring the right thing? Maybe the data are not adequate given the theory, but is the theory adequate given the data? These are hard issues to tackle, but in my review I will indicate that the author is missing a couple of opportunities where some of these more fundamental questions could have been asked and in part responded to.

The overall framework of the thesis

While I understand that the thesis consists of three unconnected pieces, it would have been helpful if the author had included a somewhat longer introduction than the present three pages. These pages now do not go much beyond summarizing the main conclusions from the paper. It would have been illuminating to learn more about the author’s view on the importance of measurement for economic analysis. One approach could have been to frame issues in the light of the increased need for harmonized statistics in a globalizing world, going beyond the conventions of the national accounts, which have been falling behind the developments in producer and consumer theory in the past two or three decades. Another approach might have been to frame the papers in the light of a political theory of measurement: how are measures used in economic policy? Why are major changes in measurement methods a bigger problem (creating an economic bias) for policy makers than measures that are biased but unchanged (creating a statistical bias). How far behind are the economic statistics in picking up issue concerning utility and welfare, rather than production and wealth?

Without addressing one or more of those underlying theoretical issues in some detail, this work can easily be characterized as “measurement without theory”, and empirical and statistics-driven economists (are which there are few anyway) are not paying themselves a service by unnecessarily being accused of that.

### Comments on chapter 1

This chapter is the most original and interesting of the three chapters in the book. It takes a serious look at three international datasets that are intensively used for economic analysis worldwide. In particular it focuses on measures of economic performance, such as GDP, GDP per capita and GDP per worker. After reviewing the datasets separately, and paying specific attention to issues such as PPP adjustments, it proceeds with a direct comparison of how strongly these variables are related between the three datasets and then tests the results from a couple of key studies on the sensitivity of the actual dataset chosen. An important conclusion is that, despite differences in results – which obviously exist – a key issue for researchers to consider is to select a dataset on the basis of the purpose of the analysis. In this light the discussion on p. 17 at the end of section 2 could be a little more extensive.

This study is very well done, and has a clear and good research design. What could have been emphasized a bit more strongly in the beginning of the chapter, is that the discussion restricts itself to comparing the measures between the datasets, but does not get into the question about the quality of the data themselves. While the basic data used in the datasets may not have differed that much, some of the additional work done on these data may have improved or worsened the overall quality. The issue also does not address issues like the use of GDI vs. GDP, market prices vs. basic prices, etc.. Also an assessment of the quality of the data for countries at different income levels would have been useful.

The issue of PPPs, index numbers (including the Gerschenkron effect and the spurious correlation effects) are not very well explained. They could have benefited from a more rigorous treatment by using some formal expressions based on index number theory. I also think the discussion on the reasons why the Gerschenkron effect does not play up could have been discussed in a little more detail than the quick observations on p. 20. In particular, the impact of service sector growth which is obviously affecting the Gerschenkron impact deserves more attention.

My biggest issue with the chapter, however, arises from the testing of the various studies in section 4. The only thing that is tested for are the performance measures to be explained, but the study does not consider the impact of alternative explanatory variables on the right hand side. These variables will often have come from the same data source, and the consistency with the variable to be explained might have been bigger in the original study than in the test. Things get even more problematic when – in the case of the Collins and Bosworth study – the numerator (GDP) of the variable is changed while the denominator (employment) is left unchanged. This significantly reduces the validity of many of the conclusions on differences in results, etc..

Finally, on the conclusions I have some trouble with the wording in the last para on “wrong choice of the dependent variable”. Here it is important to make a distinction between “wrong” on the depending on the database chosen (the purpose of this paper) and “wrong” on the basis of the type of variable (an entirely different topic).

#### Comments on chapter 2

Chapter 2 deals with the important issue of multifactor productivity growth. It goes in great detail into the measurement of a “sophisticated” growth accounts which delivers an MFP measure which is stripped for the obvious sources of variation across countries that should not really be part of the MFP residual.

This is useful and carefully done work, building on the state of the art methodology set out by OECD (2001). The chapter is a bit weak on the data issues itself (in fact the data section 3 is less than half a page), while it is here that we should have gotten a better understanding of how reliable the data inputs in the growth accounts are.

The largest contribution of the paper comes from section 4, which presents a range of sensitivity tests looking at which measurement issues most affect the MFP residual. Still I think the paper fails to discuss some of the crucial assumptions in the growth accounts framework, such as scale issue, the effects of unmeasured inputs, capacity utilization and adjustment costs. These issues ave been widely addressed in the literature (most recently see, for example, Basu and Fernald, 2001; Basu et al., 2004). It would have been helpful if the author had thought creatively how she could have used complementary data, for example, from micro datasets, on getting a better handle on some of those issues. For example, table 2 could be interpreted as a capacity utilization effect.

The impact of the relatively small but volatile effect from labour composition in Section 3 makes one wonder about whether the growth accounts are really picking up all effects from changes in labor composition. If education is seen as so important by many scholars, policy makers, etc., why are the effects so limited? And why so volatile? Labour compisition cannot move that erratically after all. A more detailed discussion of these issues whould have been welcome.

Finally, how do we interpret the MFP measure, which after all is a residual? Should we treat this as technology, market regulation effects, adjustment costs or whatever? Or is it still a “measure of our ignorance” as Abramovitz named it a long time ago? How should we interpret negative MFP (also an issue in chapter 3 of the thesis). A discussion on these issues would make this chapter less sterile and more informative for the reader.

The title of table 1 is misleading: we are not looking at an “effect” (which suggests a relationship) but at “impact” of measurement procedures.

#### Comments on chapter 3

Chapter 3 covers a more mundane issue compared to the previous two chapters, focusing on the measurement of capital services in the national accounts of the Czech Republic. Although I believe this is important work, I think the paper could have moved beyond an

excellent description of the actual measurement method, and gone into a bit more detail about the underlying assumptions, models and theory.

Importantly, the chapter does not go into issue concerning gross vs. net output (when using capital services, as developed from the productive stock, is gross output the correct output measure) and on the interpretation of asset lives (how to define an asset life? How should we measure it?). What is the effect of greater asset detail, as it increases the difference between the stock and services measures? How to deal with issues concerning the lease of capital vs. purchases?

Also the discussion on the endogenous vs. exogenous measures of rate of return is a bit meagre. What is theoretically preferable? How to interpret negative rates of return which emerge from some of the ex post measures, for example, for construction.

In sum, this chapter is great as a methodological background paper to the capital measures of the Czech Statistical Office. But for an academic piece in a dissertation it should have gone beyond the description of the data. A missed opportunity.

#### Conclusion

We need more work on the “economics of measurement”, and this dissertation is a great contribution and more than worthy of a doctorate degree. The academic profession also has the duty to guide the statisticians in not only telling them how to do things, but also providing a firm theoretical basis and analytical framework, that help to devise a statistical agenda that gets at the major economic questions before us. The candidate has the potential to contribute to this field, and I am looking forward to see more good work from her in the future.