PROFESSOR SCHNEIDER'S SHADOW ECONOMY (SSE): WHAT DO WE REALLY KNOW? A REJOINDER[1]

Edgar L. Feige[2]

Abstract

Professor Schneider's "Comment" on my "Reflections" paper fails to answer the key critiques leveled against his MIMIC estimates of the shadow economy. His "Comment" only serves to reinforce the contention that his documentation is inadequate, that his normalization procedures are arbitrary and conceptually flawed, and that his explanation of the "serious calibration error" that reversed the trend of SSE earlier estimates is untenable. A recent attempt to reproduce his findings also concludes "that it is not possible to replicate Schneider's MIMIC indexes based on the documentation from the paper." In short, the evidence challenging the veracity of his MIMIC estimates of the worldwide shadow economy is so strong as to question their place in the academic, policy and popular literature.

INTRODUCTION

Let me begin by commending Professor Schneider's continual efforts to focus attention on a subject dear to both of us. What may not be apparent to the general reader of this debate, unversed in the esoterica of national income accounting or MIMIC models, are the important areas of agreement that Professor Schneider and I share. We both champion the idea that unobserved economies are an important yet often neglected component of macroeconomic analysis having significant implications for governance in both developed and developing economies. We agree that non-compliance with generally accepted institutional rules has significant consequences for efficiency, equity and growth. Moreover, we both acknowledge that measuring the nature, extent and trend of these non-compliant behaviors is both notoriously difficult and essential if we are to understand the importance of the phenomenon and empirically assess its far ranging consequences. Among the consequences of non-compliance, we both have expressed concern about the erosion of the tax base, leading to the growth of government debt and/or a decline of public service provision. We share a concern about the potential distortion of our nations' information systems, the inequities caused by its distributional effects and the consequences of corruption and illegal activities engendering the erosion of the moral fabric of society. Finally, we have both acknowledged the strengths and weaknesses of various measurement approaches. What then are our areas of disagreement?

As much as we agree on the importance of studying this phenomenon, we approach the problem of defining it very differently, and our taxonomic differences affect our assessment of appropriate methods of measurement and of our evaluation of the state of our current empirical knowledge. I contend that there is not one "shadow economy" but many unobserved economies and set forth a taxonomic framework, which identifies their complex interrelationships. Different empirical methodologies are required to estimate the composition, magnitude and trend of these different unobserved economies. I believe that the institutions

[1] Citation: Journal of Tax Administration, 2016, Vol. 2, No.2.

[2] Professor of Economics Emeritus, University of Wisconsin-Madison. [elfeige@wisc.edu]. I wish to acknowledge the thoughtful comments of W. Lee Hansen and the invaluable econometric insights generously offered by Trevor Breusch, whose dogged commitment to seeking the truth is exemplary.

whose rules are being violated have the greatest incentive, resources, information and expertise to measure both the extent of non-compliance, and its most salient consequences. Tax revenue authorities are best suited to measure the unreported economy and its resulting tax gaps, and national and international statistical agencies are most adept at obtaining exhaustive measures of national income and output. Building public confidence in these complex measures requires extensive and timely reporting, greater transparency, greater methodological consistency between countries and over time, and the inclusion of error ranges to reflect associated uncertainty[3]. Professor Schneider's (2016) "Comment" in no way contradicts these conclusions.

However, in contrast to my taxonomic approach, Professor Schneider makes no distinctions between the various unobserved economies, preferring to employ the catch-all phrase "shadow economy" to describe them all. After various iterations, he has now settled on a "narrow" definition that identifies Schneider's Shadow Economy (SSE) solely as the "underground" economy component of the non-observed economy (NOE). Since the NOE is composed of the "underground", "informal" and "illegal" economies, it follows that the ratio of SSE/NOE <1. Regardless of whether one accepts the figures presented in my Table 2 (Feige, 2016, p.20) or Schneider's (2016, Section 2.5, Tables 2 and 3), it is clear that the values of SSE typically exceed the estimates of officially measured NOE obtained from named representatives of national statistical agencies, by several orders of magnitude. Although Professor Schneider expresses amazement[4] that the experienced statistical agencies of the Netherlands, Norway, Sweden and Turkey report such "remarkably low values" for their NOE's, his astonishment does not make their findings any less accurate.

We have both written extensively about the merits and shortcomings of various approaches to measuring unobserved activity. My own writings have decried the proliferation of Type 1 errors in the empirical economics literature (Feige, 1975) and questioned the reliability and robustness of early IRS audits and national accounting discrepancy estimates (Feige, 1989a). I described shortcomings of the Tanzi currency demand approach (Feige, 1986) and abandoning my own transactions method, I examined the implications of relaxing the assumptions of the simple currency ratio method in order to make them more realistic (Feige, 1989a). Feige and Urban (2003, 2008) demonstrated that various versions of the electric consumption method yielded unreasonable estimates of unrecorded income for transition countries. I have attempted to improve the accuracy of the general currency ratio approach with independent measures of currency velocity (Feige, 1989b) and with improved estimates of the amounts and locations of US dollars circulating abroad (Feige, 1996, 1997, 2003, 2012a, 2012b). After more than three decades of commitment to this area of research, I reluctantly felt compelled to conclude:

> Given the shortcomings of conventional macro model estimates of the underground economy and the lack of transparency and consistency of NOE estimates, it is high time that the profession acknowledges how little we really know about the underground economies and their causes and consequences. (Feige & Urban, 2008, p.287)

The "Reflections" paper (Feige, 2016) reviewed and evaluated the subsequent progress made and the remaining challenges confronting national and international statistical agencies in their efforts to improve measures of the non-observed economy (NOE). I also reviewed the

[3] Feige (2016, p.28)
[4] Schneider (2016, Section 2.5).

MIMIC/CD method employed by Professor Schneider to estimate SSE for 162 countries for the period 1999-2007. Citing the critiques of prominent economists, I concluded that "SSE estimates suffer from conceptual flaws, apparent manipulation of the results and insufficient documentation for replication, questioning their place in the academic, policy and popular literature" (Feige, 2016, p.1).

Disregarding the incisive critiques of his MIMIC/CD methodology, Professor Schneider continues to champion its results "as the most reasonable estimates of the size of the shadow economy" (Williams & Schneider, 2016, p.36). The critical question this Rejoinder must address is whether Professor Schneider's (2016) "Comments" provide any informative answers to the key challenges concerning the veracity of his work.

Specifically, does his "Comment" adequately address the charge that his documentation is insufficient and often inaccurate? Does he refute the conclusion that the meaning of the latent variable in his MIMIC model is so obscure as to question its relationship to any unobserved economy? Does he discredit the results demonstrating that his shadow economy estimates are multiples larger than they are expected to be? Does his "Comment" adequately explain the nature of his "serious calibration error" that forced him to reverse the trend and changed the size of all his worldwide shadow economy results?[5] Does he explain the implications of finding that his MIMIC index is negative and the consequences of his arbitrary decision to add a constant term to make the index positive?[6] As will be elaborated below, his oft-unsupported assertions and his inadequate and inaccurate responses only serve to reinforce my conclusions concerning the lack of veracity of his results. I shall document these charges, focusing on the major areas of concern: documentation, normalization, calibration and replication.

THE DOCUMENTATION ISSUES

Professor Schneider's findings are controversial for a variety of reasons. As the "Reflections" paper documents, there is considerable uncertainty as to whether the MIMIC methodology employed is even capable of measuring the "narrow" concept Professor Schneider now defines as his Shadow Economy. Professor Schneider acknowledges how "notoriously difficult" it is to measure "economic activity that is deliberately hidden", yet he claims to have succeeded in doing so worldwide for 162 countries for the period 1999-2007 to a reported accuracy of one decimal place. In his "Comment", he characterizes the foregoing statement as an "attack" on him (Schneider, 2016, Section 2.2) rather than the simple statement of fact it represents. He goes on to contend, without documentation: "*I **always** state that these point estimates have a margin of error of +/- 15%*"(emphasis added).[7]

A more accurate statement by Professor Schneider might have read:

> In Schneider & Williams (2013), I first announced, on page 30, but cited no evidence or documentation, that "*Estimates of the size of the shadow economy by MIMIC methods are generally thought to have a margin of error of +/-15 percent*." On page 50, I reiterated that unsupported claim stating: "*As noted, the MIMIC estimates have an error margin of +/–15.0 per cent of their estimated value*."

[5] Schneider Buehn and Montenegro (2010b, p.1)
[6] Schneider Buehn and Montenegro (2010b, p.18, footnote 24)
[7] Feige (2016, p.18, footnote 24) cites the first and only mention of the margin of error in Professor Schneider's published work.

When I discovered these surprising assertions, which I had never before encountered in Professor Schneider's papers, I wrote to him as follows:

> Since I cannot find any calculation of these error margins explicitly in your book, can you explain to me how you calculated these error margins? Have you published these error margins in any previous publication that I may be unaware of? (Email-November 21, 2013)

To date, Professor Schneider has not answered this inquiry. However, his "Comment" introduces yet another confusing element concerning the magnitudes of SSE results. When confronted with the fact that the ratio of SSE/NOE is expected to be less than one for all countries, yet turns out to be uniformly considerably greater than one, Professor Schneider discovers a rationale for suddenly reducing all SSE estimates by 20%.[8] Citing an obscure survey pertaining exclusively to Germany (Feld & Larsen, 2012, p.61), Schneider suggests that "one can make the assumption that roughly 20% needs to be deducted from these macro shadow economy measures to allow for legally-bought material which is already counted in official GDP." (Schneider, 2016, Section 2.5). The cited survey for Germany suggests that the figure could be as high as 25% and that illegal activities constituting another 27-30% are similarly included. Why not deduct 50% or 55% from the shadow economy measures? In addition, if these adjustments to SSE are reasonable now, then why has he never before applied them to his other published SSE estimates? These are just two added examples of Schneider's disconcertingly cavalier approach to documentation and data handling.

However, to comprehend fully the importance of "documentation issues" for assessing the credibility of Professor Schneider's findings, it is crucial to have a broad understanding of how SSE results are produced. SSE results depend on a two-step procedure. First, a MIMIC model is estimated, yielding a time dependent index ($\tilde{\eta}_{it}$) of the latent variable for each country. The index essentially determines the temporal trend of SSE. A second step is required to produce an estimate of the size of SSE (as a percent of recorded GDP) for a particular country at any moment in time (η^*_{it}). The MIMIC index ($\tilde{\eta}_{it}$) must now be scaled ("calibrated") to an exogenous measure of SSE for that particular country at a particular moment in time. Let ($\eta^*_{i\,2000}$) denote an exogenous estimate of SSE for country i in the year t= 2000. Then:

$$1)\quad \eta^*_{it} = \tilde{\eta}_{it} / \tilde{\eta}_{i\,2000} \times \eta^*_{i\,2000}$$ [9]

Professor Schneider claims to obtain these $\eta^*_{i\,2000}$ exogenous values from currency demand models for each of the 162 countries.[10] The importance of documenting the provenance of each these 162 exogenous values becomes apparent when one recognizes how substantially these exogenous values influence Schneider's results. Recall that the MIMIC index essentially determines the temporal trend of SSE (within country variation) while the between country variation in SSE is due to the 162 exogenous values presumably derived from independent currency demand studies. A simple analysis of variance (ANOVA) applied to SSE temporal cross-country results suggests that the MIMIC procedure accounts for less than 2 percent of the total variance of Schneider's results with more than 98 percent of the variance due to the

[8] See Schneider (2016, Tables 2 and 3, Section 2.5).

[9] This is the country equivalent of the calibration equation (7) in Schneider, Buehn and Montenegro (2010a, p.19; 2010b, p. 18) and equation (2) in 2010c, p.453.

[10] For a number of developing countries, Schneider calibrates his index to "base values originating from the year 2005 because of data availability" (Schneider, Buehn & Montenegro, 2010b, p. 18, footnote 24).

exogenous currency demand values. It seems that his much-vaunted MIMIC procedure contributes virtually nothing to the overall variation in SSE published figures.[11] It is clear that one cannot assess the veracity of Schneider's findings without investigating the provenance and reliability of the critical exogenous values that account for most of the variation in SSE results.

My requests to Professor Schneider for necessary data documentation date back to the year 2002 when he first presented DYMIMIC estimates for transition and OECD countries.[12] Since my earlier efforts were unsuccessful, I again requested documentation for the sources of his exogenous currency demand calibration values in 2008 and again in 2011, this time to Professor Buehn who initially expressed a willingness to provide me with the information after I met him at a professional conference. My data requests (repeatedly copied to Professor Schneider) included the following language:

> I would be happy to have you simply answer the following question in detail:
> **How is each of the 162 benchmark values for the 162 countries specifically derived? (**Emphasis original**)**
>
> Your paper gives the impression that these estimates come from standard Tanzi type currency demand models that are described in Appendix 1 (p.37) but this is never clearly spelled out or referenced. This issue is crucial since these benchmark values establish most of the variation across countries and many of the substantive results of your paper rely heavily on cross-country variations. (Email July, 5, 2011)

As acknowledged in my "Reflections" paper (Feige, 2016, p.18 footnote 26) a year later, (July 16, 2012) Professor Buehn provided me with the raw data for the specifications listed in their (2010a) paper. However, there was no information concerning the sources of the key benchmark (calibration) values. After repeated requests to Professor Schneider, on January 2, 2013, I received his "Preliminary Documentation of the Size of the Shadow Economy in 171 Countries" [13] whose introduction stated:

> In this preliminary documentation the calibration figures of the size of the shadow economies in 171 countries when using the MIMIC approach and their exact source (literature reference) are shown, so that everyone can check, what figures have been used. For all these "starting" figures the exact sources are given in the literature review, so that everybody can trace them further back.

Unfortunately, Professor Schneider's 165-page document did not contain a single source identifying the original specific currency demand model from which his "starting" values were derived. More than 50 percent of the referenced studies were self-referential; whose source information provided nothing more enlightening than his standard notation, "Own

[11] This result is itself surprising and may be related to the mysterious manner in which "the MIMIC index has been adjusted to the positive range by adding a positive constant" (Schneider, Buehn & Montenegro, 2010b, p.18, footnote 24).

[12] The sources for his results were "Own calculations using DYMIMIC method" (Schneider 2002, Table 2, p. 7) and "Currency demand approach, own calculations" (Schneider, 2002, Table 3, p. 13). My unsuccessful efforts to obtain data and model specifications were documented in Feige and Urban (2008, p.288, footnote 1).

[13] Professor Schneider's "Comment" (Schneider 2016, Section 3) incorrectly claims that on June 7, 2013 he sent me documentation for 179 countries. On that date, I did receive his revised documentation for 27 countries as indicated below.

calculations." Most references were to papers written after Schneider had first published his calibration values and hence could not have been their source. Upon pointing out the uselessness of this document to Professor Schneider and requesting specific references identifying original sources for his year 2000 "calibration" values for just three or five countries, on June 7, 2013, I received a document entitled, "A Preliminary Documentation of the Size of the Shadow Economy in 27 Selected Countries" which explained:

> This documentation has the purpose to provide the values of the shadow economy (in % of GDP) for 27 countries, which "served" as starting values of the calibration procedure for the MIMIC estimations of these countries, e.g. in the study by Schneider, Buehn and Montenegro (2010).

Once again, this "documentation" proved to be completely inadequate. In numerous cases, the referenced "sources" of the starting values referred to work published years after the values had already appeared in Schneider's papers. For example, the "calibration" value for Cameroon initially appeared in Schneider and Klinglmair (2004, Table 7.1 p.41) yet its "source" was listed as Suslov and Ageeva (2009). In other instances, a currency demand model was not the source of the calibration value, nor did the cited "source" value bear any correspondence to the actual exogenous value Schneider used for calibration.[14]

On November 22, 2013, Professor Montenegro provided me with his raw data for 152 countries. My subsequent attempts to reproduce these data from original sources revealed numerous discrepancies as did my attempt to reconcile the Buehn and Montenegro datasets with one another. I tabulated these discrepancies in Excel files that I sent to Buehn, Schneider and Montenegro on November 26, 2013, with further questions concerning how the data were standardized. I never received any further clarifications concerning these data discrepancies or any replies to my requests concerning how the input data had been transformed. Professor Schneider's "Comment" refers to the last email he sent me on November 25, 2013, and completely misquotes the text he sent me.[15] I naturally ignored Professor Schneider's disingenuous request for documentation from an unspecified paper instead of answering my inquiries to him.

The foregoing examples of inconsistencies, inaccuracies and ambiguities represent only a small sample of the numerous problems encountered in efforts to obtain information from Professor Schneider. They strengthen the conclusion that his lack of documentation concerning the provenance of the key "calibration" values, which explain virtually all of the variation of his results, casts a shadow on the veracity of his findings. Additional key gaps in the documentary record required for adequate replication pertain to data sources and transformations involving differencing, missing data, standardization procedures and choice of time periods. These gaps have precluded reproduction and replication of his findings for a decade.

---

[14] For example, the "source" given for Serbia was Christie and Holzner (2004) whose reported estimate for the year 2001 was 19 percent. A "Household Income Tax Method" derived this figure. Schneider's starting calibration value for Serbia was not 19 percent but incongruously 36.4 percent as published in Schneider (2007, Table 3.2.4, p. 19). Although Schneider's "documentation" specifically referred to the Schneider, Buehn and Montenegro (2010c) study, that paper does not contain any shadow economy estimate for Serbia.

[15] Schneider (2016, Section 3) states: "In a mail from November 25, 2013, I asked him ..."Would you be willing to send me similar detailed documentation of the econometric estimates of your last paper**, I mean the paper Feige & Urban published in 2008,** in order to estimate the size and development of the shadow economy?" The bolded words did not appear in his email to me. In their place were the words "**using your approach**". I had no idea to which paper or which approach he was referring.

THE NORMALIZATION ISSUE

My "Reflections" paper called attention to the consequences of Professor Schneider's choice of normalizing indicators and normalizing coefficients. Professor Schneider acknowledges that his arbitrary choice of the sign of the normalizing coefficient determines the sign of the structural parameters of the causal variables, (Dell'Anno & Schneider, 2006, p. 5) and that he chooses a (-1) normalizing coefficient on GDP in order to obtain his desired result, namely, that tax rates and SSE are positively related. I have already argued that economic theory suggests the relationship to be either ambiguous or negative (Feige, 2016, p. 22). Having guaranteed his favored result with the arbitrary choice of the (-1) coefficient on GDP or average working time, Schneider goes on to misleadingly conclude with respect to his direct and indirect tax variables, "that both causal variables are highly statistically significant and have the expected positive sign in all equations." (Williams & Schneider, 2016, p.81). His results have the expected sign because he forced them to have the expected sign, not because his data supported his hypothesis.

In his "Comment", Professor Schneider also justifies his assumption that the normalizing coefficient on GDP should be (-1) because "an increase in shadow activity has a negative effect on official GDP development" (Schneider 2016, Section 4). Yet in Schneider (2009, p.1106), he concludes that the shadow economy and official GDP are complementary, requiring that his normalization coefficient should be chosen as (+1).[16] Moreover, SSE and GDP must be positively related to one another to the extent that national income accountants capture the underground economy in recorded GDP.

THE CALIBRATION ERROR

My "Reflections" paper focused attention on the "serious calibration error" (Schneider, Buehn & Montenegro, 2010b, p.1) that forced the authors to report completely different estimates of SSE for each of the 162 countries, reversing the trend of the shadow economy from their initial findings presented in Schneider, Buehn and Montenegro (2010a). Whereas their initial (2010a) paper reported a worldwide increase in SSE, the subsequent versions of the paper revealed a mirror image reversal in the worldwide trend. To be precise, let SSE_a (s,i,t) represent Professor Schneider's shadow economy result published in the 2010a paper for specification s, country i in year t and let SSE_b (s,i,t) represent the new SSE result published in 2010b. The following relationship holds for all s, i, and t:

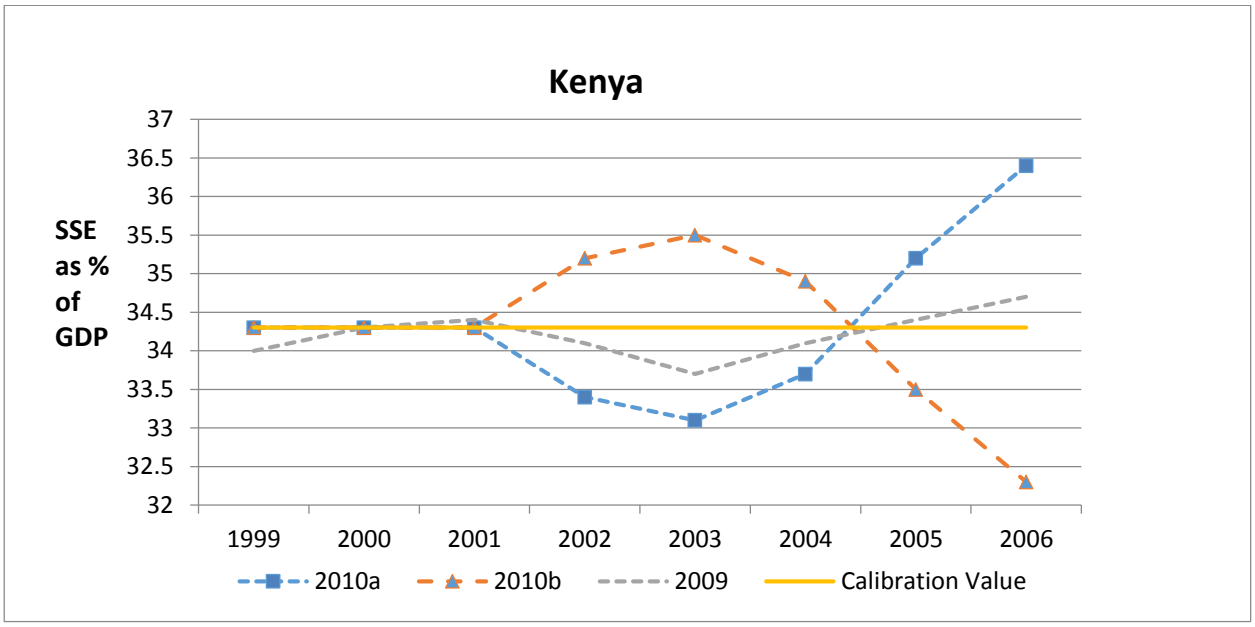
1) $\{SSE_a(s.i.t) + SSE_b(s,i,t)\}/2$ = Schneider's exogenously determined estimate of SSE for country i in year t=2000.

Thus, regardless of which specification they employ, the two curves representing the temporal path of SSE for every country are mirror images reflected about a horizontal line whose height is exactly the exogenous (calibration) estimate for the year 2000. Figure 1 displays the typical results using the example of Kenya, with the blue (square) line representing the initial faulty temporal path of SSE displayed in Schneider, Buehn and Montenegro (2010a), and the red (triangle) line representing the temporal path in Schneider, Buehn and Montenegro (2010b) after the "calibration error" was corrected.

[16] Schneider (2009, p.1106) states, "government may not have a great interest to reduce the shadow economy due to the fact that: income earned in the shadow economy increases the standard of living of at least 1/3 of the working population, and between 40 and 50 % of the shadow economic activities have a complementary character, which means that additional value added is created, and this increases the official GDP"

Figure 1

Schneider's Shadow Economy Estimates for Kenya



The (green) dashed line displays the SSE estimates reported in Schneider and Buehn (2009). Since the 2009 results display a temporal path similar to the results of the faulty Schneider, Buehn and Montenegro (2010a) study, could the 2009 study also be subject to a similar "calibration error"?[17]

The author's explanation of their mistake appeared in the opening footnote of the 2010b version that read:

> Unfortunately the estimates of the original version (WPS 5356) needed to be revised due to a serious calibration error (sign switch). We apologize for this, especially as we now have in this version a negative trend for the size and development of the shadow economies over 1999 - 2007, which we did not have in the original version. (Schneider, Buehn & Montenegro, 2010b, p.1)

Their paper contained no further explanation concerning the nature of the calibration error. Professor Schneider's "Comment" now informs us that:

> Unintentionally, a sign error occurred in an Excel file. Unemployment was shrinking for almost all countries during the years 1999 to 2007 and, due to the mistake, the positive coefficient of unemployment was multiplied by -1. Hence, we found an increase in the shadow economy. Unfortunately, this mistake, which we did not notice immediately, occurred. Realizing the mistake, we immediately corrected it, updated the dataset and published a revised version of the (working) paper. (Schneider, 2016, Section 4)

[17] These (2009) results were published in the Economics –Open Access-Open Assessment E-Journal that lists the Schneider and Buehn (2009) paper as its "most downloaded" article, having accumulated 26,396 downloads.

Empirical mistakes are readily made, and when speedily acknowledged, explained and corrected, cause little harm. However, the sign change on the unemployment coefficient that Professor Schneider now cites as their "mistake" could not have been responsible for the remarkable reversal of all of SSE results. Recall that the text in the original Schneider, Buehn and Montenegro (2010a), and the "corrected" Schneider, Buehn and Montenegro (2010b, 2010c) papers identically read:

> the MIMIC model index of the shadow economies is calculated using the structural equation (1), i.e. by multiplying the coefficients of the significant causal variables with the respective time series. For the numerical example of specification 1 the structural equation is given as:
>
> $$\tilde{\eta}_t = 0.14\ x_{1t} - 0.06\ x_{2t} - 0.05\ x_{3t} - 0.27\ x_{4t} \quad (6)$$
>
> where, x_{1t} equals the size of government, x_{2t} and x_{3t} denote the business and fiscal freedom index, and x_{4t} represents GDP per capita." [Schneider, Buehn & Montenegro (2010a, p18-19); (2010 b, p.17); (2010c, p. 453)][18]

Note that the unemployment rate was omitted from their equation (6) since its coefficient was zero and statistically insignificant in the model for 98 developing nations.[19] Since the unemployment variable was not included in the calculation of the MIMIC index, its supposed sign change could not have affected the size or trend of SSE results. Yet as illustrated by Figure 1, the authors systematically reversed SSE results for Kenya (and for all the other 97 countries) between the two versions of the paper.[20] This reversal of results occurred despite the fact that the coefficient estimates of the structural equation (6) remained the same and that the unemployment variable (the presumed source of the mistake) never entered into the calculation of the MIMIC index in either version of the paper.

How then could an error referring to the sign of a variable that was not involved in calculating the MIMIC index affect the size and trend of SSE results for all the countries in precisely the same symmetric fashion? Clearly, Professor Schneider's explanation is completely implausible. However, if the "sign switch" on the unemployment variable is not the cause of the "serious calibration error", then what was the nature of their "mistake", how was it discovered and how was it remedied? Could a clue to the dramatic reversal of the SSE results, be contained in the mysterious addition to the footnote in the revised version, which stated that "the MIMIC index has been adjusted to the positive range by adding a positive constant" (Schneider, Buehn & Montenegro 2010b, p.18, footnote 24; 2010c, p.453, footnote 8).

What did Professor Schneider's results look like before he arbitrarily added this mysterious constant to make his MIMIC index positive? Why did he find it necessary to add the constant?

---

[18] Between versions, the authors have inexplicably renumbered all the specifications in the Tables. In the original (Schneider, Buehn & Montenegro 2010a, p.17), Specification 1 presents the results for "98 Developing Countries" whereas in the "corrected" (SBM 2010b, p. 16) version, Specification 1 refers to the results for "88 developed countries." The "corrected" version (2010b) is in error since their equation (6) shows the parameters obtained for the "98 Developing Countries." Schneider, Buehn and Montenegro (2010c) reverts to the same specification numbering as appears in the original version (SBM 2010a).

[19] Schneider, Buehn and Montenegro (2010b, Table 3.1,p.16)

[20] Compare the results listed in Schneider, Buehn and Montenegro (2010a. Table 3.3.1, p. 20) with Schneider, Buehn and Montenegro (2010b, Table 3.3.2, p.21).

How large was the chosen constant and does its magnitude affect the size and trend of SSE results? Could the calibration error also have affected Schneider's earlier published shadow economy results, many of which were used in the derivative studies cited in Feige (2016, p.25-26)? These questions require answers if we are to understand the nature and full implications of the still unexplained "serious calibration error."

THE REPLICATION ISSUE

Reproduction, replication and robustness testing in economics should be quite straightforward when data sources, raw data, data transformations and statistical procedures are all fully documented and readily available. Although Professor Schneider presented MIMIC/CD results to the profession in 2002, to date, despite numerous attempts, only Trevor Breusch succeeded to replicate two of the early studies. His successful replications were the result of persistent detective work undertaken without assistance from Professor Schneider. Upon discovering the actual transformations and "benchmarking" procedures undertaken to produce the results, Breusch dismissed Schneider's complex applications of the MIMIC method, finding SSE results untenable.

My own efforts to reproduce the raw data from listed data sources were unsuccessful, as were my requests to obtain necessary clarifications from the authors. The most recent replication effort, of which I am aware, is a careful study undertaken by Marie-Astrid Maenhout (2016) who attempted to reproduce the raw data and replicate the derivation of the MIMIC index for specification (6) (25 High Income OECD countries) of the Schneider, Buehn & Montenegro (2010b) study.[21] By limiting her focus to a single specification comprising the countries with the most readily accessible data and to the replication of the MIMIC index rather than the derivation of the exogenous "calibration" figures, Maenhout increased her chances for a successful replication. Because she was unable to reproduce the raw data for three of the eight variables employed in Specification (6), she proceeded with the replication effort using the raw data that Professor Buehn had supplied to me. Following the procedures outlined in the 2010b paper, she found that her parameter estimates of the causal model had the same signs as the published estimates except for the tax burden, which was significantly negative, suggesting that higher tax burdens were associated with smaller shadow economies. The most important variable in the Schneider, Buehn and Montenegro (2010b) paper, business freedom was the least important in her attempted replication. The least important variable in the (2010b) paper, the total tax variable had the second largest impact in the replication, albeit with the opposite sign.

In order to arrive at estimates of the size of SSE, Maenhout adopted Professor Schneider's exogenous calibration values of the shadow economy for the year 2000, and benchmarked the MIMIC index she had derived as described in Schneider, Buehn and Montenegro (2010b,

[21] Ms. Maenhout's supervisor Professor Adriaenssens, on reading my "Reflections" paper, contacted me and requested the raw data that Buehn had sent to me. Adriaenssens' email read, "As we are only replicating the MIMIC index, and not the benchmarking procedure, we would benefit a lot from the raw data. We asked Prof. Schneider repeatedly, but all he managed to provide us are the final estimates. That is why we ask your help: could you provide us with these raw data? (Email –February16, 2016) quoted with the permission of Professor Adriaenssens (University of Leuven). On February 26, 2015 I sent Professor Adriaenssens the requested raw data as well as the Excel discrepancy worksheets I had prepared in attempting to reproduce and reconcile the Buehn and Montenegro datasets. After further requests, Schneider finally forwarded Buehn's data to Leuven. Ms. Maenhout's thesis summary concluded, "From the data reconstruction exercise I learned that it is not possible to reproduce the dataset based on the documentation from the paper."

equation 7). Whereas the Schneider, Buehn and Montenegro (2010b) estimates of SSE typically declined by less than two percentage points between 1999 and 2006, Maenhout's estimates fluctuated wildly with seven countries showing negative shadow economies for the year 2006. The 2006 estimates of SSE for Australia and Canada were -242% and -257% respectively. Maenhout (2016) concluded, "it is not possible to replicate Schneider's MIMIC indexes based on the documentation from the paper."

CONCLUSIONS

As much as we concur on the importance of studying unobserved economies, Professor Schneider and I approach the subject matter differently with regard to taxonomic and measurement issues. I prefer the analytic clarity of a taxonomic rule that results in definitional distinctions that correspond to empirically observable categories, whereas Professor Schneider employs the term "shadow economy" as a "catch-all". Professor Schneider interprets my "Reflections" paper as a personal "attack" on him, rather than the critical evaluation of both national accounting and macro methodologies it was intended to be. While I am personally sorry if my critiques of his MIMIC/CD method hurt his feelings, scientific accountability and informed public policy demand standards of documentation and replicability that transcend the feelings of any one individual.

Reproduction and replication are the gold standards of scientific inquiry. Replication serves to root out false claims and enables the profession to distinguish between "constructs", that is, results influenced by arbitrary decisions which bend the conclusions toward a researcher's prior opinions, and "estimates", namely, data-determined inferential outcomes obtained by applying accepted statistical procedures to coherently specified models. Not surprisingly, replication efforts are unlikely to be successful if the provenance of data sources are obscure, and the exact procedures followed in an analysis are inadequately documented.

My "Reflections" paper cited trenchant critiques of Professor Schneider's MIMIM/CD methodology, pointing out conceptual errors, non-robust results, undocumented and questionable data transformations, and concerns that arbitrary choices could substantially influence outcomes. As this "Rejoinder" documents, Professor Schneider's "Comment" is of no help in deflecting these concerns because it fails to resolve critical issues concerning documentation, normalization, calibration and hence, replication of his widely disseminated results. The most recent careful effort to replicate those results concludes, "it is not possible to replicate Schneider's MIMIC indexes based on the documentation from his paper." (Maenhout, 2016)

Professor Schneider's final offer that he and I write a joint paper "pointing to the differences between the various methods" used to measure the shadow economy fails to address the critical aforementioned issues. Over the past three decades, Professor Schneider has written so many papers and chapters repetitively describing the advantages and disadvantages of the various methods that I have stopped counting them.[22] I doubt that the profession will benefit from yet another one, and therefore I must decline his magnanimous invitation for collaboration.

[22] See for example: Alm, Martinez-Vazquez and Schneider (2004, Appendix A); Bajada and Schneider(2005, p.381-390); Schneider (1986, p. 645-649); Schneider (2005, Appendix A); Schneider, 2007, Appendix 1); Schneider (2009, p. 1114-1116); Schneider (2015, p 8-13); Schneider and Buehn (2009, Appendix A); Schneider and Buehn (2016, p 9-24); Schneider and Enste (2000, p. 91-99); Schneider and Enste (2002, Chapter 3); Schneider and Enste (2013, Chapter 3); Schneider and Williams (2013, p.27-31); Williams and Schneider (2016, Chapter 2).

Moreover, according to CollEc, Professor Schneider already holds the title of the world's top ranked economist on the scale of co-authorship centrality.[23] With eighty-five co-authors, he surely will not miss me as the eighty-sixth.

It is high time to move beyond recitations of the strengths and weaknesses of different approaches. If we are to determine which studies are replicable, robust and reasonable and which are undeserving of professional acceptance, we must penetrate procedural complexity to expose the intricate details of how results are attained. Only then can we assess whether the findings are legitimate data driven "estimates" resulting from commonly accepted inferential procedures. All the evidence to date challenges the veracity of Professor Schneider's worldwide shadow economy results. The inadequacies of his "Comment" to address this evidence only serves to reinforce the conclusions of my "Reflections" paper namely, that his findings suffer from conceptual flaws, apparent manipulation of results and insufficient documentation for replication, questioning their place in the academic, policy and popular literature.

[23] http://collec.repec.org/rank/betweenness/1.html

REFERENCES

Alm, J., Martinez-Vazquez, J. & Schneider, F. (2004). "Sizing" the Problem of Hard to Tax. *AYSPS Conference: The Hard-to-Tax: And International Perspective*, Georgia State University. 1-72.

Bajada, C. & Schneider, F. (2005). The Shadow Economies of the Asia-Pacific. Pacific Economic Review, *10*(3), 379-401.

Christie, E., & Holzner, M. (2004). Household Tax Compliance and the Shadow Economy in Central and Southeastern Europe. *wiiw Balkan Observatory*, Working Papers 038.

Dell'Anno, R. & Schneider, F. (2006). Estimating the Underground Economy using MIMIC Models: A Response to T. Breusch's Critique (Working Paper No.0607) Linz, Austria: Department of Economics, Johannes Kepler University of Linz.

Feige, E.L. (1975). The Consequences of Journal Editorial Policies and a Suggestion for Revision". *Journal of Political Economy*, *83*(6), 1291-95.

Feige, E.L. (1986). A Re-examination of the "Underground Economy" in the United States: A Comment on Tanzi". *International Monetary Fund Staff Papers*, *33*(4), 768-781.

Feige, E. L. (1989a). *The Underground Economies: Tax Evasion and Information Distortion.* Cambridge: Cambridge University Press.

Feige, E. L. (1989b). Currency Velocity and Cash Payments in the US Economy: The Currency Enigma, *Munich Personal RePEc Archive Paper 13807*. Retrieved from http://mpra.ub.uni-muenchen.de/13807/.

Feige, E.L. (1986). A Re-examination of the "Underground Economy" in the United States: A Comment on Tanzi. *International Monetary Fund Staff Papers*, *33*(4), 768-781.

Feige, E.L. (1996). Overseas Holdings of U.S. Currency and the Underground Economy. In S. Pozo (Ed.), *Exploring the Underground Economy* (pp. 5-62). Kalamazoo, Michigan: W.E. Upjohn Institute for Employment.

Feige, E.L., (1997). Revised Estimates of the Size of the U.S. Underground Economy: The Implications of US Currency Held Abroad. In O. Lippert & M. Walker (Eds.), *The Underground Economy: Global Evidence of its Size and Impact* (pp. 151-208). Vancouver, BC: The Fraser Institute.

Feige, E.L. (2003). The Dynamics of Currency Substitution, Asset Substitution and de facto Dollarization and Euroization in Transition Countries. *Comparative Economic Studies*, *45*(3), 358-383.

Feige, E.L. (2012a). New Estimates of U.S. Currency Abroad, the Domestic Money Supply and the Unreported Economy. *Crime, Law and Social Change*, *57*(3), 239-263.

Feige, E.L., (2012b). The Myth of the "Cashless Society: How Much of America's Currency is Overseas?" In *Deutsche Bundesbank Eurosystem, International Cash Conference 2012: The Usage, Costs and Benefits of Cash: Theory and Evidence from Macro and Micro Data*, (pp.119-151). Frankfort am Main*:* Deutsche Bundesbank.

Feige, E.L., (2016). Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the "Shadow Economy"? *Journal of Tax Administration*, *2*(1), 5-41.

Feige, E. L., & Urban, I. (2003). Estimating the Size and Growth of Unrecorded Economic Activity in Transition Countries: A Re-evaluation of Electric Consumption Method

Estimates and their Implications. *William Davidson Institute Working Paper Number 636*. Retrieved from https://core.ac.uk/download/pdf/3102894.pdf.

Feige, E. L., & Urban, I. (2008). Measuring underground (unobserved, non-observed, unrecorded) economies in transition countries: Can we trust GDP? *Journal of Comparative Economics*, *36*(2), 287-306.

Feld, L.P. & Larsen, C. (2012). *Undeclared work, Deterrence and Social Norms: The case of Germany*. Berlin/Heidelberg: Springer-Verlag.

Maenhout, M.A. (2016). The Underground Economy: Evaluation and Replication of Schneider's Estimates. *Dissertation Summary, University of Leuven* (KU Leuven), Brussels.

Schneider, F. (1986). Estimating the Size of the Danish Shadow Economy using the Currency Demand Approach: An Attempt. *Scandinavian Journal of Economics, 88*(4), 643-668.

Schneider, F. (2002). The Size and Development of the Shadow Economies of 22 Transition and 21 OECD Countries. *IZA DP* No. 514.

Schneider, F. (2005). Shadow economies around the world: What do we really know? *European Journal of Political Economy*, 21, 598-642.

Schneider, F. (2007). Shadow economies and corruption all over the world: New estimates for 145 countries. *Economics eJournal*, 2007-9.

Schneider, F. (2009). Size and development of the shadow economy in Germany, Austria and other OECD countries. *Revue Économique*, *60*(5), 1079-1116.

Schneider, F. (2015). Schattenwirtschaft und Schattenarbeitsmarkt: Die Entwicklungen der vergangenen 20 Jahre. *Perspektiven der Wirtschaftspolitik* 16(1), 3-25.

Schneider, F. (2016). Comment on Feige's paper "Reflections on the Meaning and Measurement of Unobserved Economies: What do we really know about the "Shadow Economy"? *Journal of Tax Administration*, *2*(2).

Schneider, F., & Enste, D. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature, 38,* 77-114.

Schneider, F., & Enste, D. (2013). *The Shadow Economy: An International Survey* (second edition). New York: Cambridge University Press.

Schneider, F., & Buehn, A. (2009). Shadow economies and corruption all over the world: Revised estimates for 120 Countries. *Economics eJournal*, Version 2, 27 October.

Schneider, F., Buehn, A., & Montenegro, C. E. (2010a). *Shadow Economies All over the World: New Estimates for 162 Countries from 1999 to 2007*. (World Bank Policy Research Working Paper 5356). Washington DC: The World Bank.

Schneider, F., Buehn, A., & Montenegro, C. E. (2010b). *Shadow Economies All over the World: New Estimates for 162 Countries from 1999 to 2007 (Revised Version)* (World Bank Policy Research Working Paper 5356). Washington DC. : The World Bank.

Schneider, F., Buehn, A., & Montenegro, C. E. (2010c). New Estimates for the Shadow Economies all over the World. *International Economic Journal*, *24*(4), 443-461.

Schneider, F., & Buehn, A. (2016). Estimating the Size of the Shadow Economy: Methods, Problems and Open Questions. (*IZA DP* No.9820).

Schneider, F., & Enste, D. (2000). Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature,* XXXVIII, 77-114.

Schneider, F. & Klinglmair, R. (2004). Shadow Economies around the World: What Do We Know? (*IZA DP* No. 1043).

Schneider, F. and Williams, C. (2013). The Shadow Economy. (IEA, Institute of Economic Affairs). London, UK: The Institute of Economic Affairs.

Suslov, N. & Ageeva, S. (2009). Influence of Energy Prices on the Size of the Shadow Economy: A Cross Country Analysis. (Project no-R06-1251, Working Paper No. 09/05E) Stockholm, Sweden: Economics Education and Research Consortium (EERC*).*

Williams, C. & Schneider, F. (2016). *Measuring the Global Shadow Economy: The Prevalence of Informal Work and Labour*. Cheltenham (UK): Edward Elgar Publishing.