APPENDIX S2 for

"Following the blind? Database Coding Policies and the Case of IFRS Noncompliance"

This supplementary appendix compares our paper with a related concurrent paper, "Deviations from the mandatory adoption of IFRS in Europe? Why non-adoption does not mean non-compliance," by Nobes, C., & Stadler, C. (2024). *European Accounting Review*, *33*(4), 1497–1519, hereafter referred to as NS.

NS, like us, express doubts about the noncompliance rates in Pownall, G., & Wieczynska, M. (2018). "Deviations from the mandatory adoption of IFRS in the European Union: Implementation, enforcement, incentives, and compliance." *Contemporary Accounting Research*, *35*(2), 1029–1066, hereafter referred to as PW. NS also document that noncompliance with IFRS adoption does not exist in the EU.

In Section A, we outline our incremental contributions beyond NS. Section B analyzes the explanations identified by NS's paper and compares these results with ours.

A. Incremental contributions beyond NS

i) Highlighting the implications of the lack of disclosure of database coding policies

A key element in our paper is highlighting the lack of disclosure of commercial databases' coding policies: i) Coding conventions that are missing in the data item's definition, ii) coding conventions that are not updated when changing the coding policy, and iii) nonindicative naming conventions of data items that suggest a different coding policy for users. We demonstrate that the lack of disclosure of the database coding policy for the *Consolidation* item has ultimately led to misperceptions about IFRS noncompliance rates.

This issue, which explains most of the PW results, is not detected by NS. Instead, NS say that errors exist in Worldscope. For that reason, while NS pinpoint a problem, they do not offer a sufficient analysis of the roots of the problem or a feasible solution going forward.

ii) Database comparison and recommendations

While NS focus on Worldscope only and call for future research on other databases, we address this directly by comparing all major data providers (Refinitiv Company Fundamental, Worldscope, Compustat Global, and Orbis) in terms of their coding policies and the accuracy of relevant data items.

Therefore, we can show that the data coding policy issue prevails among all major providers (i.e., not limited to the Refinitiv Company Fundamental data used by PW or the Worldscope data used by NS for the *Consolidation* variable). As the case illustrates, *none* of the data providers offers a feasible solution, and we need other workable solutions in case commercial databases fail.

iii) Showcasing that extracting accounting data from primary sources can be done cost-effectively by utilizing state-of-the-art natural language processing (NLP) models

Our database comparison documents that commercial data providers may not offer a feasible solution for certain accounting data items. We therefore showcase the feasibility of extracting raw data from corporate filings for international firms using NLP models with high accuracy rates. We show that today, even for non-U.S. firms (without EDGAR, based on PDF filings of reports, e.g., from the Perfect Information database), basic accounting data crucial to the inference of a study

can be extracted in an automated, large-sample-based manner. This insight matters for archival studies where commercial databases do not offer appropriately extracted data or with insufficient coverage of the accounting data item.

We further suggest that future research collaboration should enhance the programming codes for data extraction by fine-tuning or increasing the quality of the codes or training data. NLP models will offer increasingly powerful and cheaper solutions compared to the time-consuming manual data collection that NS suggest as a solution to the database issues they have detected.

iv) In-depth institutional framework

Based on *primary* regulatory sources, we supplement and update the PW tables to reflect the requirements and exemptions of the EU's IFRS regulations as well as country-specific local regulations that must be considered when determining whether a firm must apply IFRS.

Specifically, we provide i) a comprehensive conceptual map of all criteria that need to be checked when determining whether a firm is mandated to adopt IFRS, ii) an updated and dynamic table on the member states additional requirements that fall outside the scope of the IFRS regulations, iii) a country-by-country description of the consolidation exemptions for member states where consolidation is a determining factor for mandating IFRS, and iv) offer guidance on the best sources to identify IFRS mandated firms empirically.

Our approach can guide future research on any EU regulation for which scope exemptions exist. For example, the CSRD in Europe offers adoption options that EU Member States must implement into national law. NS do not provide such a framework.

v) Layers of compliance

Our study has a broader scope, as we not only show that noncompliance with IFRS adoption does not exist but also rule out noncompliance with EU regulations at earlier, less visible stages. Specifically, we show that firms comply with the requirements for preparing consolidated financial statements, which trigger the IFRS mandate. NS only show that noncompliance with IFRS adoption does not exist, but they do not investigate whether noncompliance could occur at earlier stages.

B. Comparative analyses of sources of error

NS's paper and ours discuss why a firm might not be required to apply IFRS and offer alternative explanations for PW's results. However, we can better identify the underlying sources of error and their relative contribution for several reasons.

B.1. Sample and data providers

i) NS and PW noncompliance samples differ substantially

NS state that they are replicating PW's study but do not have the list of ISINs used by PW nor the $Mandatory_{it}$ classification assigned by PW. Instead, they examine three self-constructed samples comprising 1,297 non-IFRS adopting observations, i.e., entities not using IFRS according to the Worldscope item WC07536. Subsamples are (see Table S2.1 Panel A):

- 2007 sample of 403 non-IFRS adopting firms from the UK, representing an *earlier year* of IFRS adoption in the country with the largest sample size and the most alleged noncompliance cases.

- 2012 sample of 503 non-IFRS adopting firms from the UK, Germany, Austria, and Portugal. This selection includes the two largest countries in PW's sample, which also have the strongest enforcement regimes, as well as two countries with less stringent enforcement regimes in *the most recent year of PW's sample*.
- 2020 sample of 391 non-IFRS adopting firms from the UK and Germany. This year is *not* part of PW's sample period, as it is used to reflect the more recent situation.

To evaluate NS's data, we matched the PW data that we could access, which PW had kindly shared with us at the start of our project, with the NS data available on GitHub.[1] Of the total NS sample of 1,297 non-IFRS adopting observations reported in NS's Table 3, we estimate that a maximum of 394 observations are classified as noncompliant by PW.[2] Thus, a maximum of 30.4% of the NS sample overlap with PW's noncompliant sample.

While differences in sample selection are not problematic per se, they matter in this case.

First, 39.5% of NS's observations correspond to cases that PW correctly classified as nonadopters without considering them as noncompliant. Including such a high proportion of observations for which PW had already found legitimate reasons for nonadoption leads NS to overstate the extent of various data issues. This is particularly evident in NS's explanation [6], where they argue that the ESMA list was insufficient to compensate for deficiencies in Worldscope's listing information (NS, p. 1506). This concern becomes far less relevant when those valid nonadopter observations are excluded, and the focus is solely on what PW classified as noncompliant cases.

Second, including valid nonadopter cases obscures the fact that NS do not directly address the question of what went wrong. Aware that they are analyzing nonadopters in general (which can or cannot be noncompliant firms), NS accept the high frequency of their explanation [2], which states that there is no error, implicitly suggesting that this category relates to observations that PW did not classify as noncompliant. However, had the analysis focused exclusively on noncompliant firms, NS would have realized that cases for explanation [2] should not appear at all, particularly for firms in the UK and Germany, where only consolidated entities fall under the IFRS mandate. A firm not preparing consolidated statements in these jurisdictions cannot logically be classified as noncompliant—unless PW made a mistake. By treating all nonadopters as a single group, explanation [2] seems plausible. It discourages further scrutiny, which would have revealed that PW did not rely on Worldscope to determine consolidation status but instead used a different data source—an insight only visible through a precise replication of the disputed cases.

Third, given that Worldscope implemented a data coding policy change in 2013—reclassifying firms without subsidiaries that issued single-entity financial statements into "consolidated"—the impact of that change is limited to the 2012 sample in PW's dataset (as it includes some 2013 data based on their year definition) and does not affect earlier years. NS's use of 2020 as an out-of-sample benchmark is thus unsuitable for drawing inferences about the PW sample period, and NS's conclusion about the impact of the Worldscope consolidation item is overstated.

In contrast, we analyze the *same* firm-year observations that PW flagged as noncompliant (see PW's Table 2 Panel C). Thus, we cover 1,412 out of 1,538 noncompliant cases, or 91.8% of PW's

[1] See https://github.com/christian-stadler/ifrs-database/tree/main.

[2] Since NS's Excel workbook does not include the non-IFRS reporting observations from Austria and Portugal reported in Table 3 of their paper (14 observations), it is impossible to reconcile this data with PW's classification. Therefore, we conservatively assume that PW would have classified these observations as noncompliant.

sample (see our Table 2). Using our comprehensive framework, we connect these observations to secondary data sources to detect the relevant sources of errors and their relative weights.

ii) NS and PW consolidation variables are from different databases

NS use Worldscope for consolidation data, while PW use Refinitiv.[3] This difference is crucial for analyzing the source of error in PW's data.

Specifically, PW use "TF.RF.IsConsolidated" from the previous Thomson One platform to identify a firm's consolidation status. This item proves to be the primary source of error in the noncompliance sample, accounting for approximately 87% of error cases. When we compare our results based on the Refinitiv consolidation variables equivalent to the one used by PW—"TR.F.FundamentalConsolidation" and "TR.IsISConsolidated"—with those of Worldscope used by NS, we notice that most cases in the PW noncompliance sample would be correctly classified; hence, what NS capture is mainly related to observations outside of PW's sample period and does not generalize to the period before 2012.

B.2. Classification Framework

iii) Differing research questions

While NS's work and our paper document that noncompliance with IFRS adoption does not exist in the EU, the scientifically relevant follow-up question arises as to *why* the literature came to incorrect conclusions. NS focus their explanations and classification framework on "seek[ing] to explain why many firms are showing in Worldscope as not publishing IFRS statements" (NS, p. 1503). Instead, our research focus is broader in that we ask, "What went wrong?"

iv) NS framework combines sources of error with valid exemptions

NS's classification framework includes six explanations of why non-IFRS reporting could be legitimate, i.e., why a firm is still compliant, even when Worldscope codes a firm as not using IFRS:

- [1] Firm not incorporated in the country (Worldscope error)
- [2] Firm does not provide consolidated statements – Case 1: Worldscope records this (correctly)
- [3] Firm does not provide consolidated statements – Case 2: Worldscope error
- [4] Firm provides IFRS statements (but Worldscope records a different GAAP)
- [5] Firm not listed in the year (although Worldscope contains accounting data)
- [6] Firm not listed on a market that requires IFRS

Notably, these six categories represent valid "explanations why reporting is legitimate" (NS Table 3) and do not necessarily represent sources of error. For example, explanation [2] is not a source of error, as Worldscope data is correct. Therefore, when 36.9% of NS's observations are assigned to category [2] and even 72.2% of UK observations in 2007, the relevant question about the source of error in PW remains unanswered.

[3] PW confirmed that they used Refinitiv.

v) Empirical comparison of sources of error

Based on their analysis, NS argue that legitimate reasons vary by country and year, the Worldscope database contains errors in many different fields, and the most frequent error is in the item "Accounting Method For Long Term Investment>50% (WC07531)".

To validate these conclusions, in Table S2.1 Panel B, we split NS's sample (NS Table 3) into three subsamples: NS sample classified as noncompliant by PW, NS sample *not* classified as noncompliant by PW, and NS sample outside the PW sample period.

Based on the complete NS sample, one concludes:

- For 2007, explanation [2] dominates, with 72.2% of the cases, followed by [6] with 14.9%, and [3] with 7.7%.
- For 2012, explanation [2] dominates in the UK with 51.9%, followed by [3], but vice versa in Germany, Austria, and Portugal. Explanation [6] ranks third.
- For 2020, explanation [3] dominates, followed by [6].

On that basis, NS conclude that the "reasons for legitimate non-adoption vary by year and country" (see Abstract). They identify explanation [3] (Worldscope error) and explanation [6] as the most significant sources of errors.

However, when splitting the sample, results only align for the NS sample *not* classified as noncompliant by PW and the NS sample *outside* PW's sample period. A different picture emerges for the relevant NS sample classified as noncompliant by PW:

- For the UK in 2007, conclusions do not hold. Explanation [2] accounts for 91.3% of cases, indicating that Worldscope is correct and the source of error is not addressed. Explanation [3], when Worldscope is incorrect, explains less than 9%. Explanation [6] plays no role.
- Again, for the UK in 2012, explanation [2] represents 67.1% of the cases, which is not a source of error. However, explanation [3] has now become significant due to the undisclosed data coding policy change in Worldscope, which was implemented starting in 2012. Explanation [6] plays no role.
- For Germany in 2012, explanation [3] has become significant. Explanation [6] is insignificant with 2.6% of cases.

Hence, based on the 2012 sample (the last year of PW's sample period), one can see that the Worldscope variable regarding consolidation started to cause errors, which is even more drastic in the 2020 sample. This illustrates how out-of-sample data drive NS's conclusions regarding the causes of error and do not generalize to the data that PW used.

In summary, while NS's approach points in the direction of errors in PW's noncompliance data, their conclusions on sources of error are inaccurate in a replication study, where 30.1% of observations fall outside the relevant period and 39.5% do not represent errors in the replicated study. This approach leads to inflated, insignificant sources of errors (e.g., explanation [6]), overestimates unspecific Worldscope errors (e.g., due to Worldscope's later data policy change), and conflates valid exemptions with errors (e.g., explanation [2] where Worldscope is correct).

TABLE S2.1 NS Samples Explained

		Complete NS Sample							**NS sample classified as noncompliant by PW**			**NS sample *not* classified as noncompliant in PW**			**NS sample *outside* PW's sample period**	
Panel A: Firms in NS sample																
		[I]	[II]	[III]	[IV]	[V]	[VI]	[VII]	[I]	[II]	[V]	[I]	[II]	[V]	[VI]	[VII]
		DE	UK	AU	PT	UK	DE	UK	DE	UK	UK	DE	UK	UK	DE	UK
		2012	2012	2012	2012	2007	2020	2020	2012	2012	2007	2012	2012	2007	2020	2020
Total		**248**	**241**	**8**	**6**	**403**	**214**	**177**	**38**	**146**	**196**	**210**	**95**	**207**	**214**	**177**
Panel B: Explanations																
		[I]	[II]	[III]	[IV]	[V]	[VI]	[VII]	[I]	[II]	[V]	[I]	[II]	[V]	[VI]	[VII]
		DE	UK	AU	PT	UK	DE	UK	DE	UK	UK	DE	UK	UK	DE	UK
		2012	2012	2012	2012	2007	2020	2020	2012	2012	2007	2012	2012	2007	2020	2020
[1] Firm not incorporated in the country (Worldscope error)	N	-	2	-	-	16	1	4	-	-	2	-	2	14	1	4
	%	-	0.8	-	-	4	0.5	2.3	-	-	1	-	2.1	6.8	0.5	2.3
[2] Firm does not provide consolidated statements – Case 1: Worldscope records this (correctly)	N	53	125	2	-	291	8	-	14	98	179	39	27	112	8	-
	%	21.4	51.9	25	-	72.2	3.7		36.8	67.1	91.3	18.6	28.4	54.1	3.7	
[3] Firm does not provide consolidated statements – Case 2: Worldscope error	N	113	77	6	4	31	122	154	20	47	14	93	30	17	122	154
	%	45.6	32	75	66.7	7.7	57	87	52.6	32.2	7.1	44.3	31.6	8.2	57	87
[4] Firm provides IFRS statements (but Worldscope records a different GAAP)	N	6	6	-	-	1	3	8	3	1	1	3	5	-	3	8
	%	2.4	2.5	-	-	0.3	1.4	4.5	7.9	0.7	0.5	1.4	5.3	-	1.4	4.5
[5] Firm not listed in the year (although Worldscope contains accounting data)	N	4	2	-	-	4	7	2	-	-	-	4	2	4	7	2
	%	1.6	0.8	-	-	1	3.3	1.1	-	-	-	1.9	2.1	1.9	3.3	1.1
[6] Firm not listed on a market which requires IFRS	N	72	29	-	2	60	73	9	1	-	-	71	29	60	73	9
	%	29	12	-	33.3	14.9	34.1	5.1	2.6	-	-	33.8	30.5	29	34.1	5.1

Note: This table expands NS Table 3. Their sample comprises firms with NoIFRS, defined as those for which Worldscope's WC07536 is neither 'IFRS' (#23) nor 'International standards' (#02). We use NS's publicly accessible data to match it with the ISINs from PW, dividing the sample into three subsamples: "NS sample classified as noncompliant by PW," "NS sample *not* classified as noncompliant by PW," and "NS sample *outside* PW's sample period." We maintain column numbers defined by NS across all three samples to ensure consistency.