Online Appendix
Better two eyes than one: a synthesis classification of exchange rate regimes

Cécile Couharde* and Carl Grekou†.

This online appendix presents the auxiliary analyses and associated results. Section D is devoted to the exploration of the disagreement points between the two classifications. Section E regroups all supplementary or intermediate results.

Appendix D. Documenting the disagreements

D.1. Overview of the evidence

D.1.1. The de facto exchange rate regime classifications

While both the *LYS* and *RR* classifications infer exchange rate regimes based on what countries effectively do, they differ considerably regarding: *(i)* the data, *(ii)* the key statistic(s), and *(iii)* the methodology they use for categorizing the different ERR.

The *LYS* classification combines available information on the exchange rate and reserves' movements to capture the effect of interventions on the exchange rate and determine the *de facto* flexibility of ERR. On the methodology side, it builds on a cluster analysis which partitions data points (a data point corresponding to a given country's currency x at particular time t) into different ERR categories according to their similarity across the following variables: (i) *changes in the nominal exchange rate* —measured as the average of the absolute monthly percentage changes in the nominal exchange rate during a calendar year, (ii) *the volatility of these changes* —computed as the standard deviation of the monthly percentage changes in the exchange rate, and (iii) *the volatility of the net-reserves-to-the monetary base ratio*. The principle underlying this clustering is that countries experiencing low volatility of their exchange rates (in both levels and changes) and high volatility of their reserves should be associated to a *Fixed* ERR. Instead, floaters should be associated with highly volatile exchange rates (in both changes and levels) and stable reserves. By definition, intermediate regimes fall between these two extreme regimes.

*EconomiX-CNRS, University of Paris Nanterre, France. Email: cecile.couharde@parisnanterre.fr.
†CEPII and EconomiX-CNRS, University of Paris Nanterre, France. Email: carl.grekou@cepii.fr.
The synthesis classification is available here

The *RR* classification is based exclusively on exchange rate variations.[1] These are calculated as the absolute percent changes in the monthly nominal exchange rate averaged over a five-year rolling window —two-year in some cases. In addition to considering a more extended period, the *RR* classification also distinguishes itself by considering, when available, the parallel market exchange rates.

Tables D.1.1.1 and D.1.1.2 report the different regime categories identified by the *LYS* and *RR* classifications. The *LYS* classification differentiates only four categories of regimes (plus one associated with *inconclusive determinations*) that can be grouped into the usual tripartite categorization: *Fixed, Intermediate*, and *Floating*. The *RR* fine classification, by contrast, distinguishes fourteen categories that can be aggregated into five categories within a coarse classification.

Table D.1.1.1 — Levy-Yeyati & Sturzenegger *de facto* classification

Five-way classification		*Three-way classification*
Regime	Code	Regime
Inconclusive determination	5	
Free float	4	Floating ERR
Dirty float	3	Intermediate ERR
Dirty float/Crawling peg	2	
Fix	1	Fixed ERR

Note: we reverse the original *LYS* classification so that a higher category is associated with more flexibility (except regime 5).

Table D.1.1.2 — Reinhart & Rogoff *de facto* classification

Regime	Classification	
	Fine	Coarse
No separate legal tender	1	1
Pre announced peg or currency board arrangement	2	1
Pre announced horizontal band that is narrower than or equal to +/-2%	3	1
De facto peg	4	1
Pre announced crawling peg	5	2
Pre announced crawling band that is narrower than or equal to +/-2%	6	2
De facto crawling peg	7	2
De facto crawling band that is narrower than or equal to +/-2%	8	2
Pre announced crawling band that is wider than or equal to +/-2%	9	3
De facto crawling band that is narrower than or equal to +/-5%	10	3
Moving band that is narrower than or equal to +/-2% (i.e., allows for both appreciation and depreciation over time)	11	3
Managed floating	12	3
Freely floating	13	4
Freely falling	14	5

[1] The *RR* classification uses inflation data to distinguish a specific category, the "Freely falling" category. This category includes observations with a twelve-month rate of inflation above 40 percent. This category also consists of countries that have switched from a fixed or quasi-fixed regime to a managed or independently floating regime following an exchange rate crisis.

D.1.2. The disagreements between the two classifications

The above considerations suggest that the *RR* and *LYS* classifications differ substantially from a methodological perspective. The *LYS* classification appears to be a purely statistical method-based classification, while the *RR* classification benefits from *a priori* thresholds delimiting the categories. The recourse to *a priori* thresholds is a feature shared by several *de facto* classifications, which entails an undeniable methodological anchoring effect. On the other hand, the *RR* classification relies on the volatility of the parallel market exchange rates —over a five-year window— instead of the official ones —within a year— as the other classifications do. Hence, from a methodological point of view, the *RR* and *LYS* classifications appear "diametrically" opposed —with alternative classifications falling between these two extremes.

Not surprisingly, the two classifications exhibit significant divergences in the history of regimes. To illustrate this, we collapse the *RR* classification into three categories to fit the traditional three-way classification. Following the literature, we aggregate the different ERR categories of the *RR* classification as follows. The *Fixed* ERR comprises the categories 1 to 4 (fine classification), the *Intermediate* ERR includes categories 5 to 11, and the *Floating* ERR consists of the remaining categories.[2] Figure D.1.2.1 regroups the charts on the three main ERR categories' evolutions according to the *RR* and *LYS* classifications. As can be seen, for advanced economies, the *LYS* and *RR* classifications diverge substantially regarding all the categories, especially regarding the share of *Floats* and *Fixed* from 1999 onwards. The pictures regarding the *Fixed* category for the emerging and developing countries are more similar in dynamics, but the *RR* classification tends to attribute a lesser share to *Fixed* — to the benefit of *Intermediate*. With the *LYS* classification, one can again observe *Floats*' prevalence —compared to the *RR* classification.

Table D.1.2.1 presents two-way contingency tables between the *RR* and *LYS* classifications —and across different groups of countries. Considering the whole sample, the observed agreement rate between the *RR* and *LYS* classifications reaches 57.7%.[3] However, this rate differs considerably across groups of countries. It varies from around 64.4% for developing countries (DCs) to about 52.2% for emerging market economies (EMEs) and 42.8% for advanced economies (AEs). Figure D.1.2.2 maps out these disagreements. Also, as reported in Figure D.1.2.3, there is evidence of increasing discordance over time

[2]Following the literature, we exclude the "freely falling" category from the empirical analysis. This omission represents a loss of 397 observations. Furthermore, note that the "separating line" between the *Intermediate* and the *Float* ERRs is itself a source of disagreements. The selected "line" maximizes the concordance (a gain of 89 points) between the two classifications —and is in line with the literature.

[3]The agreement rate corresponds to the sum of observations along the diagonal divided by the total number of observations.

between the classifications —especially in EMEs and DCs.

On average, the agreement between the two classifications is the highest for the *Fixed* regime category, followed by the *Intermediate* category, except among AEs where this latter category presents the lowest agreement rate. These first findings are consistent with the observation made by most empirical studies that the *de facto* classifications do not overlap very well.

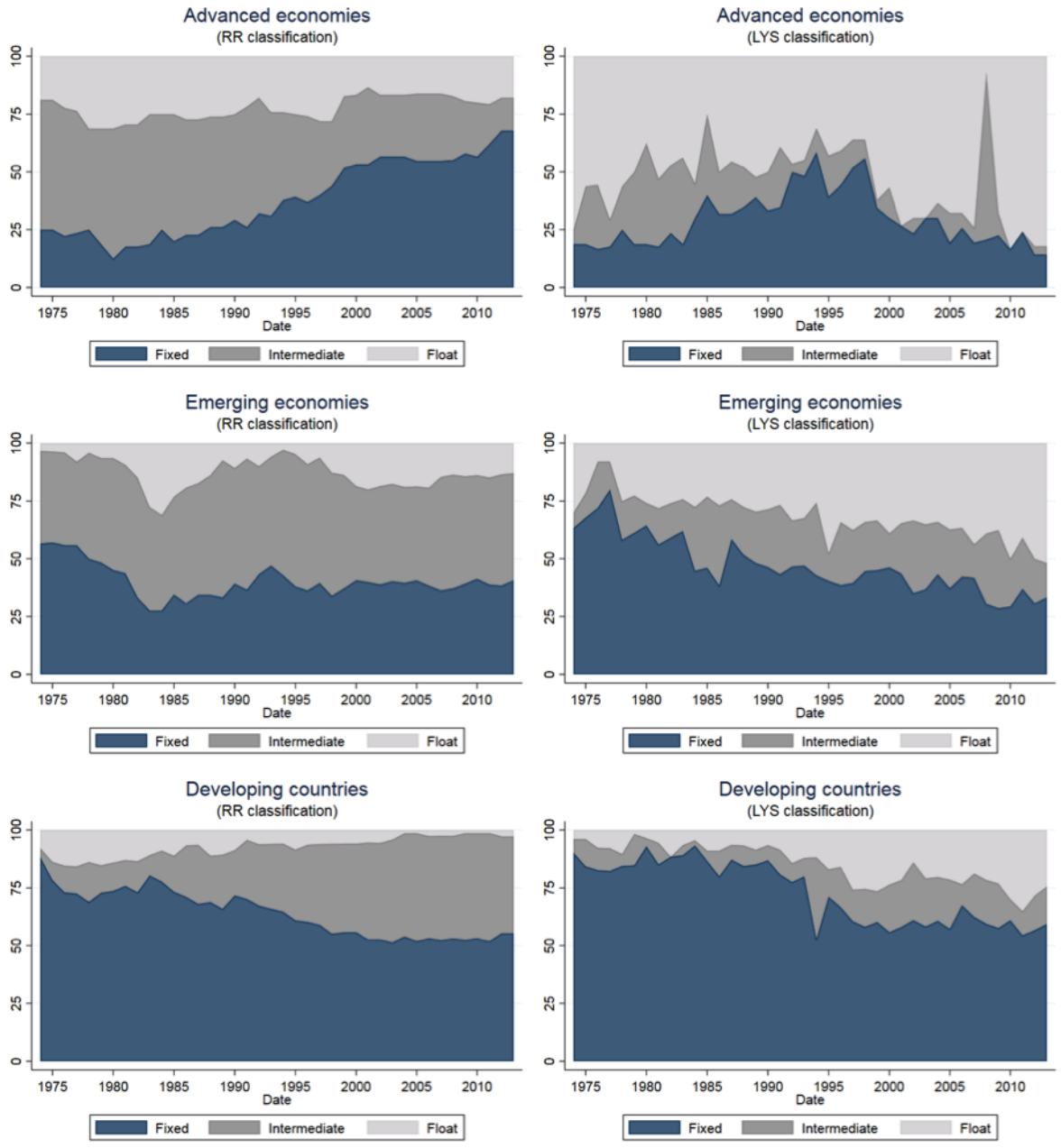

Figure D.1.2.1 — *De facto* regime distributions over time (% of annual observations)

Table D.1.2.1 — Two-way contingency tables, traditional three-way classifications

All countries

		LYS			
		Fix.	*Inter.*	*Float*	Total
RR	*Fix..*	**2080**	187	289	2556
	Inter.	481	**497**	888	1866
	Float	151	124	**314**	589
	Total	2712	808	1491	5011

Pearson $\chi^2(4) = 1.6e+03 \mid Pr = 0.000$

AEs

		LYS			
		Fix.	*Inter.*	*Float*	Total
RR	*Fix..*	**174**	36	195	405
	Inter.	101	**78**	184	363
	Float	18	27	**167**	212
	Total	293	141	546	980

Pearson $\chi^2(4) = 103.81 \mid Pr = 0.000$

EMEs

		LYS			
		Fix..	*Inter.*	*Float*	Total
RR	*Fix.*	**519**	83	41	643
	Inter.	140	**212**	399	751
	Float	54	52	**109**	215
	Total	713	347	549	1609

Pearson $\chi^2(4) = 602.24 \mid Pr = 0.000$

DCs

		LYS			
		Fix..	*Inter.*	*Float*	Total
RR	*Fix..*	**1387**	68	53	1508
	Inter.	240	**207**	305	752
	Float	79	45	**38**	162
	Total	1706	320	396	2422

Pearson $\chi^2(4) = 924.14 \mid Pr = 0.000$

Note: The different matrices represent the two-way contingency tables between the *RR* and *LYS* classifications (whole sample as well as sub-samples). Pearson $\chi^2(.)$ displays the statistics and *p*.value associated to the independence test of rows and columns –in a two-way table.

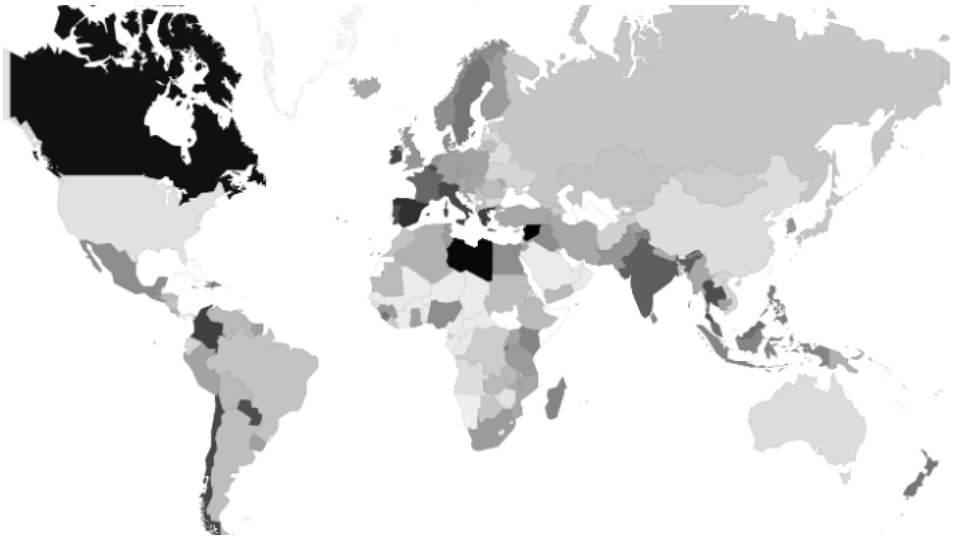

Figure D.1.2.2 — Disagreements map

Note: The shades of grey indicate the level of disagreement (number of disagreement points) between the *RR* and the *LYS* classifications (i.e. the darker, the more the disagreements)

To get a better picture of the disagreements between the two classifications, Table D.1.2.2 shows the contingency table for the whole sample of countries using the different original categories defined by the two classifications —plus additional information from the *LYS* classification. As reported above, the *LYS* classification relies on a cluster analysis —based on three variables— to determine the *de facto* ERR. In a nutshell, the algorithm (the Kinetic Monte Carlo) assigns the data to five homogeneous groups, each representing an ERR category (except the *Inconclusive determination* group). For comparability

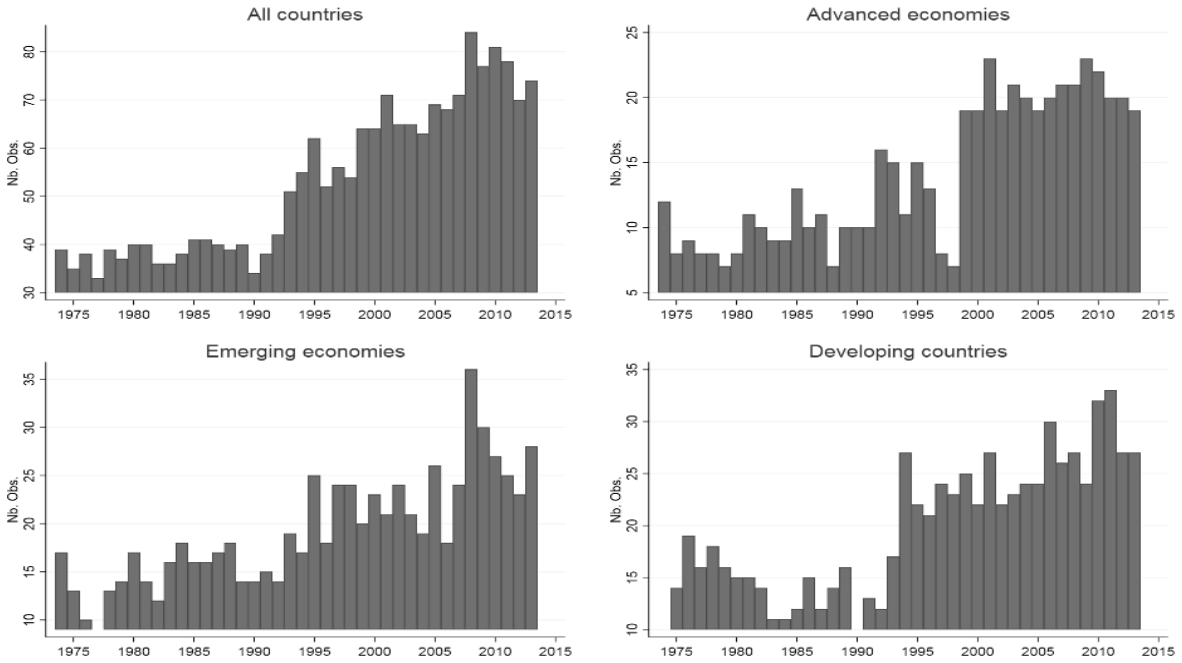

Figure D.1.2.3 — Evolution of the classification disagreements (number of observations)
Note: The height of the bars indicate the number of disagreements per year.

purposes, the two percent upper-tail of the observations for each of the three key variables are in a first step excluded. The remaining points are *z*-normalized then classified.[4] This initial stage (or first round) assigns several observations to the different ERR and leads to a considerable number of observations classified as "*Inconclusive*".[5] These latter observations are then assigned to the different ERR categories, through a second round procedure using the same methodology as in the first one. Finally, some observations left unclassified (either "*inconclusive*" or "*unclassified*") are assigned to the different ERR categories on an *ad hoc* basis using additional information. Specifically, observations so far classified as inconclusive and *(i)* exhibiting zero volatility in the nominal exchange rate, or *(ii)* considered as a *de jure* peg by the IMF with an average volatility of the nominal exchange rate smaller than 0.1%, are assigned to the *Fixed* ERR.[6]

In Table D.1.2.2, each row represents an ERR category defined by the *RR* classification. The column headers of each sub-table correspond to the data labels from the *LYS* classification procedure. "*U*" stands for "*Uncontroversial*," i.e., observations classified apart from the methodology using additional information; "*O*" indicates observations labeled as *outliers*; "*R2*" indicates observations classified in the second round and "*I*" stands for "*Inconclusive*" (more specifically *Fixed inconclusive*). The disagreements between the two classifications that are of interest to us are reported in the off-diagonal sub-tables.

[4]The excluded data points correspond to the outliers.
[5]The inconclusive category contains observations with low volatility regarding the three key variables.
[6]The same approach is used to classify the countries that have been excluded due to a lack of data.

Table D.1.2.2 — Two-way contingency table (All countries; zoom-in)

RR classification		LYS classification: Fixed					LYS classification: Intermediate					LYS classification: Float				
		Total Obs.: 2080					Total Obs.: 187					Total Obs.: 289				
		ERR	*Obs.*	*U*	*O*	*R2*	*I*	**ERR**	*Obs.*	*U*	*O*	*R2*	*I*	**ERR**	*Obs.*	*U*	*O*	*R2*	*I*
Fixed	**RR**	**1**	955	112	32	270	415	**1**	30	13	13	1	0	**1**	188	187	0	0	0
		2	862	68	5	286	372	**2**	59	0	6	36	0	**2**	41	0	0	12	0
		3	1	0	0	1	0	**3**	0	0	0	0	0	**3**	0	0	0	0	0
		4	262	23	3	78	87	**4**	98	0	0	91	0	**4**	60	0	0	23	0
	LYS							**2**	155	13	0	128	0						
								3	32	0	19	0	0						
		Total Obs.: 481					Total Obs.: 497					Total Obs.: 888				
		ERR	*Obs.*	*U*	*O*	*R2*	*I*	**ERR**	*Obs.*	*U*	*O*	*R2*	*I*	**ERR**	*Obs.*	*U*	*O*	*R2*	*I*
Interm.	**RR**	**5**	7	0	2	0	0	**5**	33	0	1	29	0	**5**	10	0	0	3	0
		6	14	0	0	1	9	**6**	17	0	1	14	0	**6**	17	0	0	8	0
		7	92	4	13	27	0	**7**	168	0	2	155	0	**7**	168	0	0	65	0
		8	134	6	7	33	24	**8**	169	0	5	126	0	**8**	416	0	0	123	0
		9	5	0	0	0	0	**9**	4	0	0	4	0	**9**	16	0	0	2	0
		10	160	13	2	14	94	**10**	75	0	6	25	0	**10**	173	0	0	26	0
		11	69	15	4	3	12	**11**	31	0	0	27	0	**11**	88	0	0	24	0
	LYS							**2**	465	0	0	380	0						
								3	32	0	15	0	0						
		Total Obs.: 151					Total Obs.: 124					Total Obs.: 314				
		ERR	*Obs.*	*U*	*O*	*R2*	*I*	**ERR**	*Obs.*	*U*	*O*	*R2*	*I*	**ERR**	*Obs.*	*U*	*O*	*R2*	*I*
Float	**RR**	**12**	147	22	7	27	38	**12**	100	0	5	21	0	**12**	189	0	0	11	0
		13	5	0	0	2	0	**13**	24	0	0	0	0	**13**	125	0	0	0	0
	LYS							**2**	103	0	0	21	0						
								3	21	0	5	0	0						

Notes: "*ERR*" stands for the regime categories specified in Table D.1.1.1 (*LYS* classification) and Table D.1.1.2 (*RR* classification). "*U*" (resp. "*O*", "*I*") stands observations labeled as *Uncontroversial* (resp. *Outliers*, *Inconclusive*); "*R2*" indicates observations classified in the second round.

D.2. The disagreements: an in-depth analysis

To gain more insights, we conduct a qualitative content analysis of each type of observation that involves a disagreement. To gain readability, these different types of disagreement are mapped out, and their evolution over time is also reported in a chart.

D.2.1. RR class. "Fixed" and LYS class. "Intermediate"

187 observations are classified as *Fixed* in the *RR* classification but *Intermediate* in the *LYS* classification. In the latter classification, 128 of these 187 observations are assigned through the second round procedure, 13 are labeled uncontroversial, and 19 are outliers. The charts below reveal both the countries and years for which the disagreements are the most significant. As can be seen, for the euro area member countries and the CFA zone countries, the divergence between the two classifications is the most important. Some eastern European and Asian countries and few Latin American countries are also concerned by this divergence. The bottom chart further shows that in 1994 and 2008, the difference between the two classifications has been most pronounced.

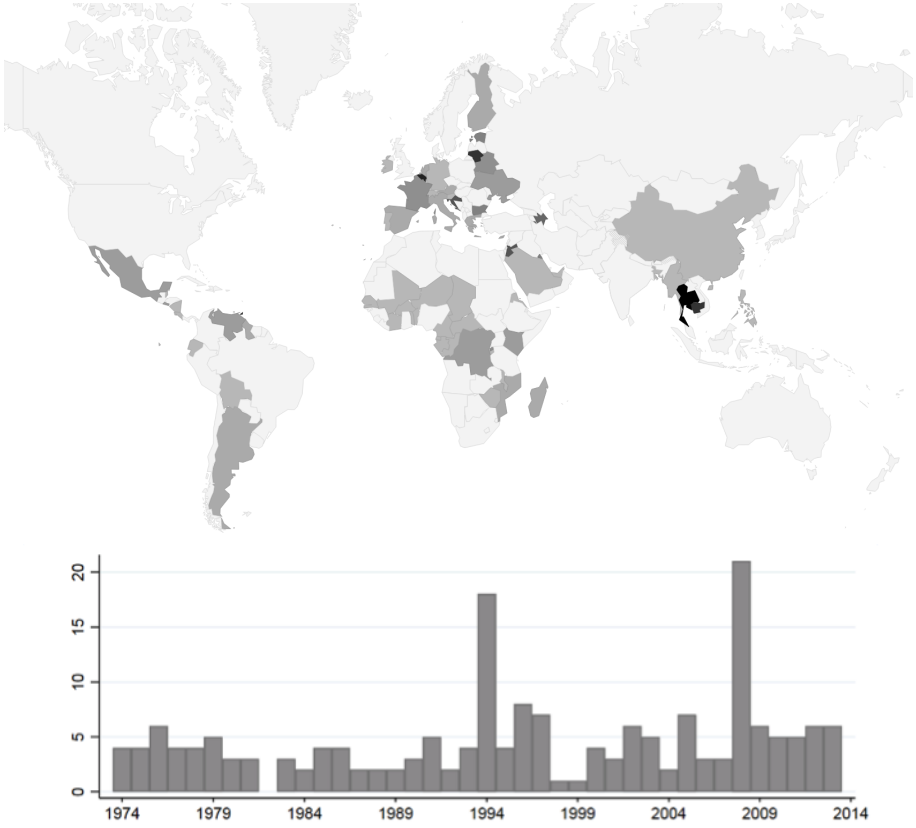

Figure D.2.1 — *RR* class. "*Fixed*" and *LYS* class. "*Intermediate*"
Notes: The map (top) displays the countries' coverage and the frequency (as reflected by the shades) of disagreements. The bottom chart displays the number of disagreements by year.

The spike in 1994 corresponds to the 100% devaluation of the CFA franc. During this

episode, ERR of the CFA zone countries were identified as *Fixed* by the *RR* classification. They were classified as *Intermediate* and as outliers (13 out of the 19 points) by the *LYS* classification.[7] The peak in 2018 is related to the ERR of some eurozone countries.[8] They have been classified as *Intermediate* during this single year and labeled uncontroversial by the *LYS* classification.

As shown in Table D.1.2.2, the remaining disagreement points correspond to 128 points classified through the second round procedure and 27 observations without any label. Regarding these latter observations, all of the 27 points correspond either to devaluation episodes (i.e., Costa Rica 1974, Kenya 1981, Philippines 1997, Ecuador 2000) or to changes in the anchor currency(ies) (i.e., Jordan 1975, Burundi 1983, Argentina 1991, Lithuania 2002). The explanation of the remaining disagreement points is less obvious. Note, however, that around a dozen disagreements (per country) are associated with few countries (i.e., Thailand 1978 and 1986-1996; Belgium 1975-76, 1978-1980, 1983-1986 and 1994).

Combining the information in Table D.1.2.2 and Figure D.2.1 for the above set of disagreement points (i.e., *Fixed* in *RR* and *Intermediate* in LYS) leads to the following observation: one of the potential sources of disagreements between the two classifications is the difference in the time horizon considered by the classifications. Indeed, adopting a year-by-year approach, the *LYS* classification does not put exchange rates' changes into a historical/broader context and therefore puts too much emphasis on the change within a year.[9]

D.2.2. RR class. "Fixed" and LYS class. "Float"

289 observations are classified as *Fixed* in the *RR* classification but *Float* in the *LYS* classification. 187 of these 289 observations are labeled uncontroversial in the *LYS* classification, while 35 are assigned through the second round procedure. As shown in Figure D.2.2, the European region, specifically the Eurozone, is the most concerned by these disagreements. In particular, 187 "*Uncontroversial*" points identified in Table D.1.2.2 involve the euro area member countries. This fact is consistent with the significant jump in the number of disagreements observed since 1999 when the euro was established —depicted

[7] The remaining 6 points labeled as outliers correspond to one-time devaluations (Rwanda 1974, Mexico 1976, Argentina 1985, Nicaragua 1991, Bulgaria 1997, Venezuela 2011).

[8] More specifically, 13 eurozone countries: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Slovenia, and Spain.

[9] However, this question of the time horizon to consider addresses the exchange rate regime definition. Should the exchange rate regime reflect only the exchange rate's behavior in a particular year, or should it instead view the change in the exchange rate in a broader context, therefore considering economic and political shocks/decisions?

on the bottom chart (Figure D.2.2).[10] Levy-Yeyati and Sturzenegger (2016) noted it was a deliberate choice to classify the ERR of eurozone member countries as *Float*, given the behavior of the euro *vis-à-vis* other currencies.[11] In the *RR* classification, the ERR is instead classified as *Fixed* –to reflect the lack of monetary policy autonomy associated with the single currency.

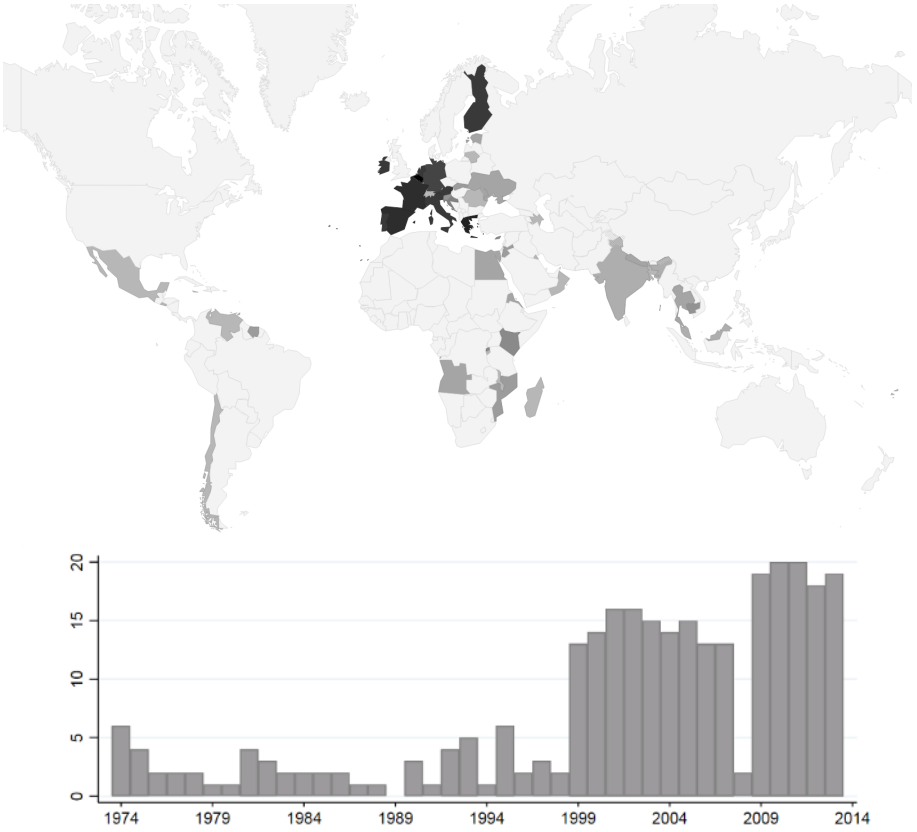

Figure D.2.2 — *RR* class. "*Fixed*" and *LYS* class. "*Float*"
Notes: The map (top) displays the countries' coverage and the frequency (as reflected by the shades) of disagreements. The bottom chart displays the number of disagreements by year.

Removing these 187 uncontroversial points associated with the euro area countries leaves 102 disagreement points with a relatively even distribution. Among these 102 disagreement points, one is classified in the *RR* category 1 (i.e., "No separate legal tender"),[12] 41 in the *RR* category 2 (i.e., "Preannounced peg or currency board arrangement"), and 60 in the *RR* category 4 (i.e., "*De facto* peg").

Most of the divergence points classified as *Fixed* –category 2– by the *RR* classification occur one year before a change in the anchor currency. Other disagreements coincide with a

[10]The 2008's fall corresponds, as noted above, to the reclassification of these countries as *Intermediate* in the *LYS* classification.
[11]The authors acknowledge that the ERR's classification for the euro area countries (i.e., *Fixed* or Float) remains an open question and that the answer depends on the issue at stake.
[12]This disagreement point corresponds to the introduction in 1997 by Eritrea of a new currency pegged to a new anchor (the US dollar).

change in the ERR within a year. They also correspond to some devaluation/reevaluation episodes (i.e., frequent devaluations in Kenya between 1982 and 1986; Maldives 2001; Venezuela 2010; reevaluation of the Nepalese rupee in 1993). The picture for the 60 observations classified as *Fixed* category 4 by the *RR* classification is less clear. However, as the difference between the two classifications involves the two extreme regimes, it is possible to come up with several explanations for these disagreements' sources. The main reason is the difference between the two classifications regarding the reference currency against which the nominal exchange rate volatility is calculated. For instance, in the *RR* classification, the Kenyan shilling volatility is measured *vis-à-vis* the SDR (Special Drawing Rights) from 1976 to 1991. The *LYS* classification uses the SDR as the reference currency over a different period (between 1975 and 1986). Another explanation is that the volatility measure and, most specifically, the definition of the threshold values delimiting the different ERR categories differ between the two classifications.

Contrary to what prevails in the *RR* classification, the *LYS* classification threshold values are determined by the algorithm/data. It follows that the same observation will be classified into two distinct regimes as long as the *LYS* procedure's threshold value will differ from that of the *RR* classification. The disagreements between the two classifications on the ERR of Belgium in 1974, 1977, 1981-82, and 1993 illustrate this point. Indeed, for this country and over a more extensive period encompassing the disagreement years, the reference currency was the same in the two classifications, changes in foreign reserves were relatively stable, the parallel market premia were negligible, and the dynamics of the official and the parallel market exchange rates were similar. Hence, the disagreement could only stem from the difference in the threshold values used by the two classifications for delimiting the ERR categories for this country.

D.2.3. RR class. "Intermediate" and LYS class. "Fixed"

481 observations are classified as *Intermediate* in the *RR* classification but as *Fixed* in the *LYS* classification. Almost 40% of these disagreements correspond to an *ad hoc* categorization in the *LYS* classification. These observations can be divided into two groups. The first group corresponds to the 139 observations labeled *inconclusives* (more precisely *Fixed inconclusives*) and assigned arbitrarily to the *Fixed* ERR category. This group consists of 18 countries, among which 5 account for more than 60% of the disagreement points: Syria (33 points between 1975 and 2011), Libya (17 points between 1988 and 2013), Egypt (14 points between 1974 and 1988), Brunei Darussalam (12 points between 1999 and 2013), and Paraguay (10 points; from 1974 to 1981 then 1987-1988). The second group corresponds to the 38 observations labeled *uncontroversial* (*Uncontrover-*

sial fix). It includes Afghanistan 2002, Brunei Darussalam 1984-98, Equatorial Guinea 1980-1984, Guinea 1976-82, Hungary 1974 and 1979, Seychelles 1978, Syria 1992-92 and 2012-13, and Vanuatu 1978-1980.[13]

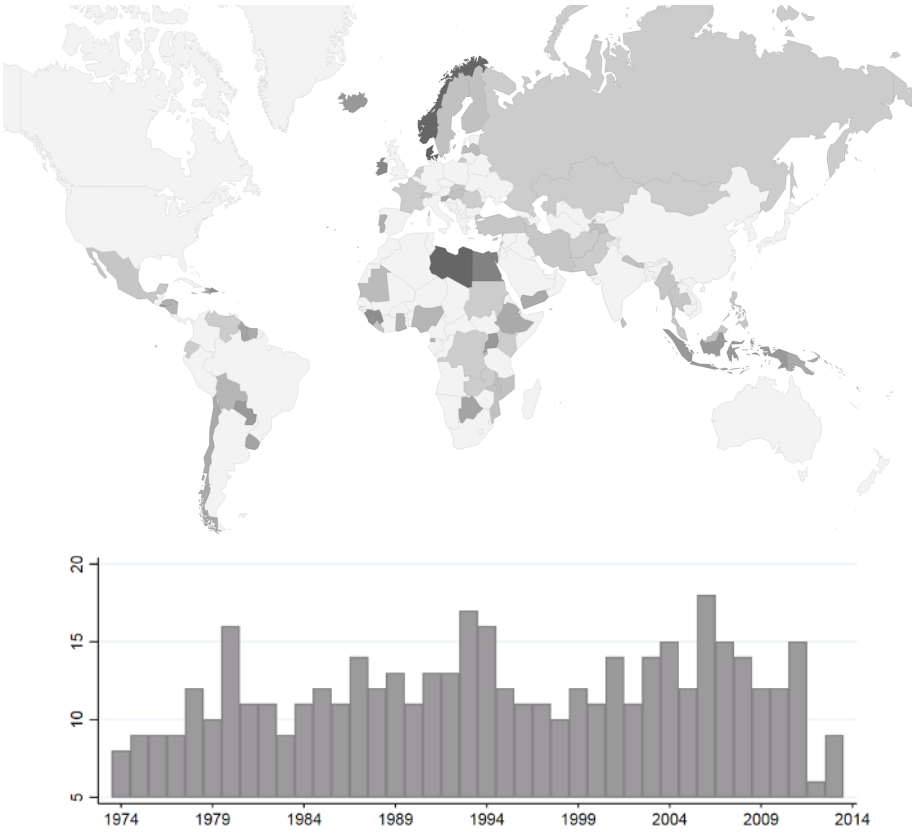

Figure D.2.3 — *RR* class. "*Intermediate*" and *LYS* class. "*Fixed*"
Notes: The map (top) displays the countries' coverage and the frequency (as reflected by the shades) of disagreements. The bottom chart displays the number of disagreements by year.

Removing these disagreement points leaves 304 observations, among which 78 are classified through the second round procedure in the *LYS* classification, and 28 are labeled as outliers.[14] As before, there is a significant concentration of the disagreement points in relatively few countries, Denmark and Norway having the highest score. Denmark (resp. Norway) is associated with 20 (resp. 19) disagreement points over the 1974-1998 period (resp. the 1992-2010 period).[15]

Intuitively, the nature of the divergence between the two classifications (Intermediate in the *RR* classification and *Fixed* in the *LYS* classification) suggests that the discriminating element(s) comes from how the two classifications assess exchange rate dynamics.[16]

[13]Again, the label "uncontroversial" refers to observations classified on an ad hoc basis because of the unavailability of classification variables.
[14]Also note that removing the points classified on an ad hoc basis considerably reduces the number of disagreement points in the *RR* category 10 (i.e., "De facto crawling that is narrower than or equal +/-5%").
[15]Ireland and Iceland also belong to this top group with respectively 13 (between 1981 and 1996) and 10 (between 1999 and 2013) disagreement points.
[16]Indeed, in the *LYS* classification, both *Fixed* and *Intermediate* regimes are associated with high reserves

However, it is more complicated than it seems because this assessment can vary across several dimensions. The first aspect to investigate is the exchange rate volatility. We first notice that 40 of the 304 remaining disagreement points coincide with a difference in the reference currency against which the exchange rate volatility is measured.[17] Furthermore, among the 264 remaining disagreement points, 120 can be associated with the parallel market exchange rate in the *RR* classification. However, among these disagreement points, few display high premia. This fact indicates that the differences between the two classifications in the definition of the threshold values delimiting the ERR categories and/or in the time horizon —over which the exchange rate volatility is measured— can also be considered as potential suspects.[18]

D.2.4. RR class. "Intermediate" and LYS class. "Float"

This fourth configuration is the one with the highest number of disagreement observations (888 points). It represents 41.7% of the total number of disagreements and covers 108 countries. As shown (Figure D.2.4), Canada and India have the highest number of disagreement points (both countries have 25 disagreement points). They are closely followed by Israel (21 points), Colombia-Malaysia-Switzerland (20 points), and Guatemala-Pakistan-Philippines-Sri Lanka-Tunisia (19 points).

Differences in the reference currency explain only 105 disagreement points. The remaining disagreement points do not present any particularity that we could use to isolate specific observations, such as observations labeled *inconclusive* and/or *uncontroversial*. However, the nature of the disagreements gives some intuitions regarding their sources. Specifically, they could be related to either the way exchange rate dynamics are assessed or the use of the official reserves, or even both. Hence, we are forced to adopt a step-by-step approach for simplicity's sake at this analysis stage. Moreover, to facilitate the analysis, we take a different approach consisting of comparing observations classified *Intermediate* by the two classifications (i.e., consensual *Intermediate* ERR) and observations classified *Float* by the *LYS* classification but *Intermediate* in the RR. Doing so allows us to compare the dynamics of the key variables for the two groups. Indeed, in the *LYS* classification, the *Float* ERR is associated with highly volatile exchange rates (both in changes and levels) and stable reserves, while the *Intermediate* ERR is associated with highly volatile reserves and exchange rates.[19]

volatility.

[17] Also, note that the reference currency's difference concerns 41 points labeled Inconclusive, 1 point labeled Outlier, 24 points classified in the second round, and 29 points labeled Uncontroversial.

[18] Only 32 (resp. 46) display premia higher than 10% (resp. 5%).

[19] Except for crawling pegs, which are associated with low volatility of the exchange rate changes.

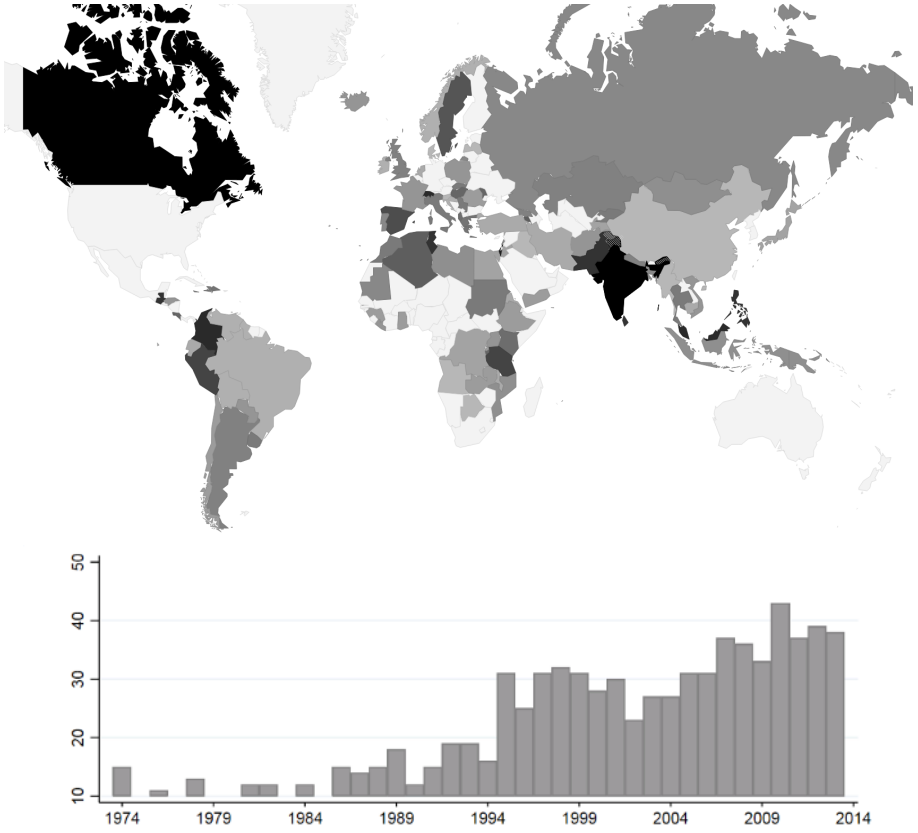

Figure D.2.4 — *RR* class. "*Intermediate*" and *LYS* class. "*Float*"
Notes: The map (top) displays the countries' coverage and the frequency (as reflected by the shades) of disagreements. The bottom chart displays the number of disagreements by year.

Noticing that the *RR* classification categories 8 and 10 (resp. "*De facto* crawling band that is narrower than or equal to +/-2%" and "*De facto* crawling band that is narrower than or equal to +/-5%") are those the more affected by the disagreements leads us to focus on the exchange rate volatility. We investigate the effect of using the parallel market premium in the *RR* classification by comparing the volatility in the changes of the official and parallel market exchange rates. Among the remaining observations, 347 are associated with the parallel market exchange rate in the *RR* classification instead of the official in the *LYS* classification. Moreover, the correlation between the volatility of parallel market exchange rates movements and that of the official one is low (0.21 on average), arguing, therefore, in favor of using the parallel market exchange rate as a possible explanation for the disagreements.[20] However, this explanation should not hide the potential role of the threshold values delimiting the ERR categories, which are higher in the *LYS* classification than in the *RR* classification. Regarding the role that might be played by the use of the official reserves' volatility, we cannot, at this stage, go further than make an assumption. This issue will be addressed further in the empirical analysis as it requires keeping all other

[20]We, however, found correlations higher than 0.90 regarding the exchange rate changes.

variables constant.

D.2.5. RR class. "Float" and LYS class. "Fixed"

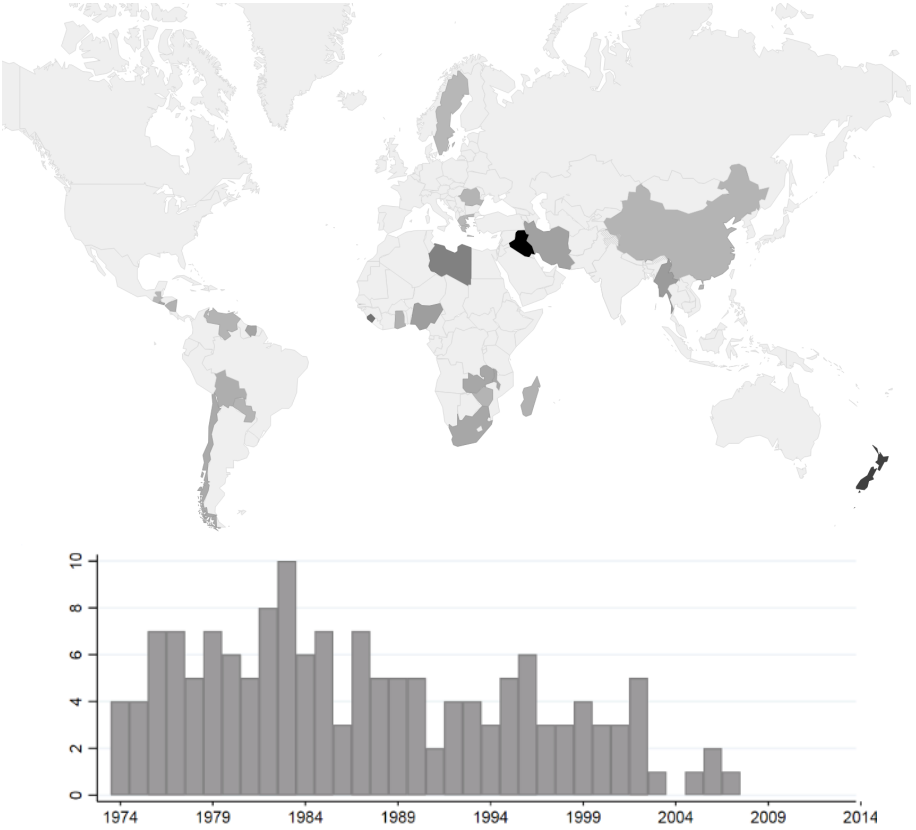

Figure D.2.5 — *RR* class. "*Float*" and *LYS* class. "*Fixed*"
Notes: The map (top) displays the countries' coverage and the frequency (as reflected by the shades) of disagreements. The bottom chart displays the number of disagreements by year.

Among the 151 disagreement points included in this fifth configuration, only 22 observations are labeled as *Uncontroversial* (China in 1987 and 1988; Iraq from 1983 to 2002), and 38 are labeled as *Fixed inconclusives* in the *LYS* classification.[21] Removing these 60 data points leaves 91 observations, among which 7 are considered outliers and 27 points are classified via the second round procedure. New Zealand is the country the more concerned by this type of disagreements (i.e., *Fixed* in *LYS* but *Float* in RR) with 16 points between 1985 and 2006, followed by Sierra Leone (13 points), Libya (10 points), and Nigeria (9 points).

New Zealand's case appears to be driven by the difference between the two classifications regarding the reference currency against which the exchange rate volatility is measured. Indeed, the *RR* classification uses as the reference currency the Australian

[21] These points correspond to: Bolivia (1975-76, 1978), El Salvador (1984-85, 1987-89), Guatemala (1987), Iran (1979, 1982-84, 90, 96, 99), Myanmar (1976, 78, 80-84, 92, 95), Paraguay (1982-1983), Suriname (1982-85, 88-90), Venezuela (1983-85).

dollar while the *LYS* classification uses the US dollar. The reference currency's difference also seems to explain the disagreement between the two classifications for Sierra Leone between 1974 and 1981 (except 1978).[22] Overall, the difference in the reference currency explains 26 disagreement points.

This leaves 66 data points, among which 41 are associated with exchange rate premia greater than or equal to 10%, with a correlation between the official and the parallel market exchange rates varying between -0.4 and 1. Hence, the use of parallel market exchange rates in the *RR* classification could also be at stake for some countries.

D.2.6. RR class. "Float" and LYS class. "Intermediate"

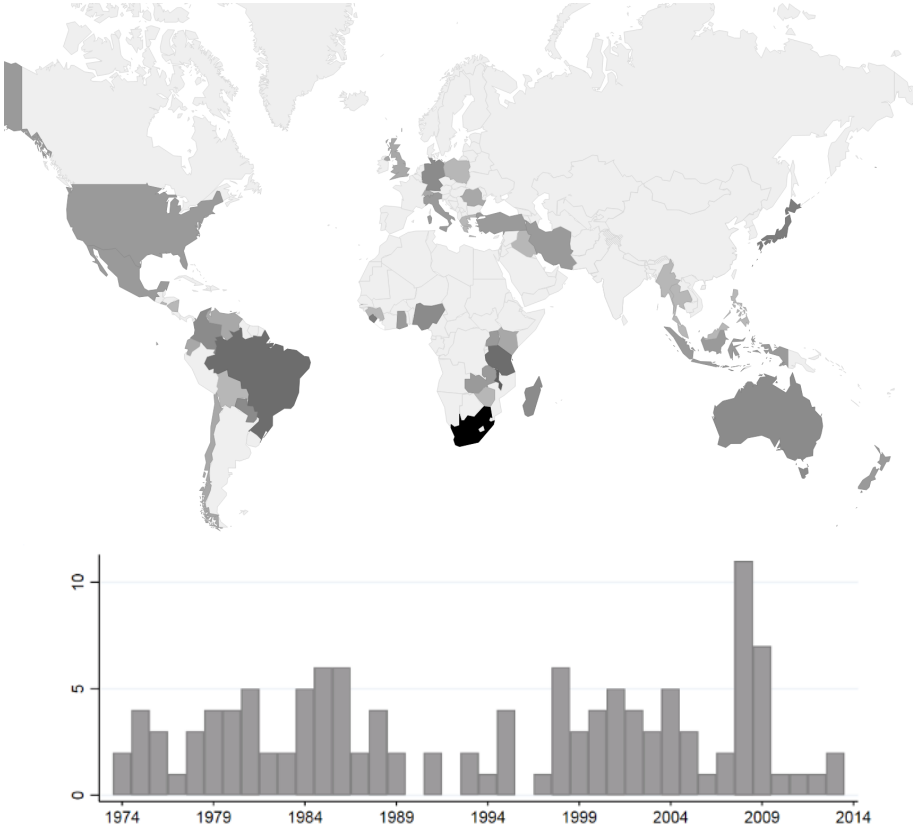

Figure D.2.6 — *RR* class. "*Float*" and *LYS* class. "*Intermediate*"
Notes: The map (top) displays the countries' coverage and the frequency (as reflected by the shades) of disagreements. The bottom chart displays the number of disagreements by year.

[22] Greece in 1982 and Myanmar in 1979 are also concerned.

As shown in Table D.1.2.2, the last type of disagreement consists of relatively few observations. Only 124 observations are classified as *Float* by the *RR* classification while classified as *Intermediate* by the *LYS* classifications. South Africa appears as the country the more concerned by this disagreement (12 points). It is followed by Malawi (7 points), Brazil, and Tanzania (6 points for each country). Among these 124 disagreement points, 38 correspond to a divergence in the reference currency against which the exchange rate volatility is measured.[23]

The remaining 86 data points share the particularity to be all classified as "Dirty float/Crawling peg" (LYS 3-way classification, category 2).[24] This category differs from the *Float* regime —in the *LYS* classification— due to the volatility of exchange rate changes and official reserves volatility. As before, disentangling each factor's effects proves to be a difficult/impossible task for such a descriptive analysis. This issue will then also be addressed in the econometric analysis.

E. Additional results

E.1. Testing the randomness of the agreements

Given that the two classifications disagree half the time, we deemed it relevant to test the concordances' randomness between the classifications. Indeed, it does not make sense to explain why the classifications diverge if the concordant observations are themselves "random". Say differently, before going any further, we must ensure that we are not seeking logic where there might be none. To do so, we compare our dependent variable —scoring 0 when the *LYS* and *RR* classifications concord, 1 otherwise— with simulated variables. More specifically, we draw respectively N —ranging from 1000 to 10000 with an increment of 1000— random dichotomous (0;1) variables of 5011 observations each time and compute for each of the simulated variables the concordance rate with our dependent variable. Figure E.1 reports the distributions of the obtained concordance rates per number of draws. As can be seen, regardless of the number of draws, the distribution appears centered around 0.5, suggesting that the simulated data only coincide —on average—with half of the dependent variable's observations. Given the considerable number of draws, one can therefore conclude that the concordance points between the *LYS* and the *RR* classifications are not random. Similar results are obtained when considering the analysis sample.

[23] These points correspond to Australia, Germany, Greece, Iran, Italy, Japan, Madagascar, Myanmar, New Zealand, Romania, Sierra Leone, Switzerland, Tanzania, the United Kingdom, the United States, Uganda, and Zambia.

[24] In the *RR* classification, there are 78 (resp. 7) points in the "Managed floating" (resp. "Freely floating") category.

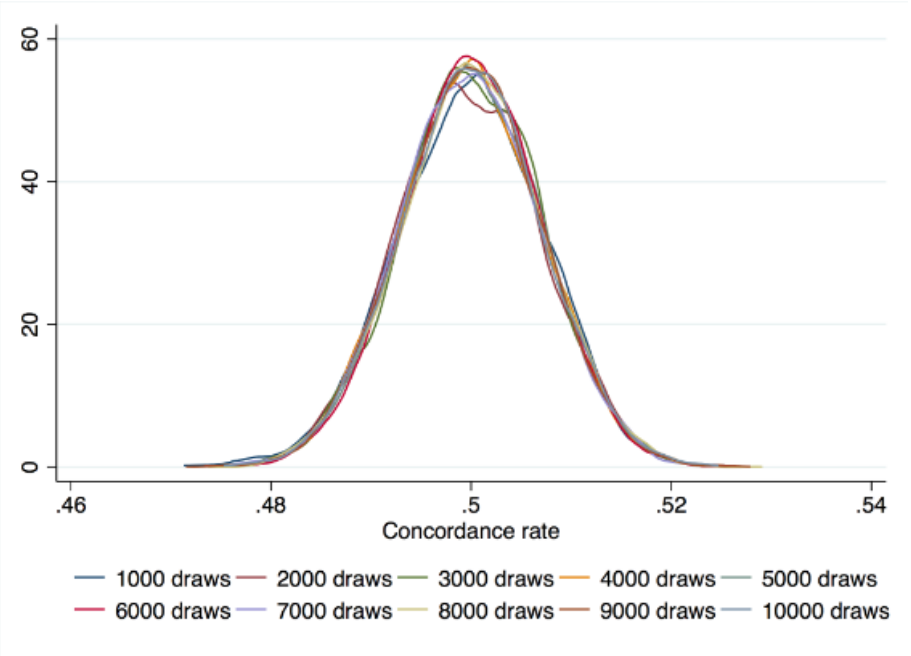

Figure E.1 — Distributions of the concordance rates with the simulated data

E.2. Coping with the multicollinearity between the LYS exchange rate volatility measures

Table E.2.1 — Principal components (eigenvalues)

Component	Eigenvalue	Difference	Proportion	Cumulative
Component 1	1.95029	1.90058	0.9751	0.9751
Component 2	0.0497084	.	0.0249	1.0000

Table E.2.2 — Principal components (eigenvectors)

Variable	Component 1	Component 2	Unexplained
Volatility of the ER	0.7071	0.7071	0
Volatility of the ER change	0.7071	-0.7071	0

E.3. Alternative estimation procedures (Floating ERR sample)

The issue we address in this appendix is related to the disequilibrium of our dependent variable's categories in the *Float* ERR sample (146 "0" and 1047 "1") and the potential associated bias. The problem is not precisely the rarity of events but rather the relatively small number of cases on the rarer of the two outcomes. We relied on penalized-Logit (Firth method) —designed for rare events— to assess whether our probit model-based estimates are biased. To enable comparison —and so to evaluate the extent of the correction in the penalized-Logit, we also reported Logit estimates.

Instead of comparing the different coefficients (more specifically converted coefficients), we follow Amemiya (1981) and focus on the probabilities. Figure E.3 plots in this regards

the different estimated probabilities of disagreements and the distributions (kernel estimates) of these probabilities. As can be seen, the simulated probabilities are similar, indicating that the disequilibrium of our dependent variable categories does not plague our Probit estimates.

Table E.3 — Logit estimations (*Floating* ERR sample)

Estimation proc.	Logit			Probit
	Robust	No correction	Penalized#	
	Betas (Std. Err.)	Betas (Std. Err.)	Betas (Std. Err.)	Betas (Std. Err.)
Horizon/Premium	-0.591*** (0.107)	-0.591*** (0.109)	-0.593*** (0.106)	-0.346*** (0.064)
E.R. volatility	-1.282 *** (0.394)	-1.281*** (0.268)	-1.217*** (0.241)	-0.773*** (0.182)
Reserves	-16.376*** (0.556)	-16.376 (386.78)	-5.105*** (1.456)	-6.237*** (0.519)
Outlier	7.469 (4.889)	7.469 (1140.6)	4.651 (3.314)	4.817** (2.264)
Round 2	2.015*** (0.472)	2.016*** (0.433)	1.974*** (0.416)	0.922*** (0.198)
Constant	18.062*** (0.593)	18.062 (386.81)	6.775*** (1.462)	7.258*** (0.538)
Pseudo R^2	0.1518	0.1518		0.1546
Log likelihood	-376.08	-376.08	-370.95	-374.81

Notes: "*Betas*" stand for standardized coefficients (except dummy variables). "***" (resp. "**" and "*") indicates statistical significance at 1% (resp. 5% and 10%). Robust standard errors are reported in parentheses. "Equiv." corresponds to the ratio between the considered Logit model and Probit coefficients. #: Firthlogit method.

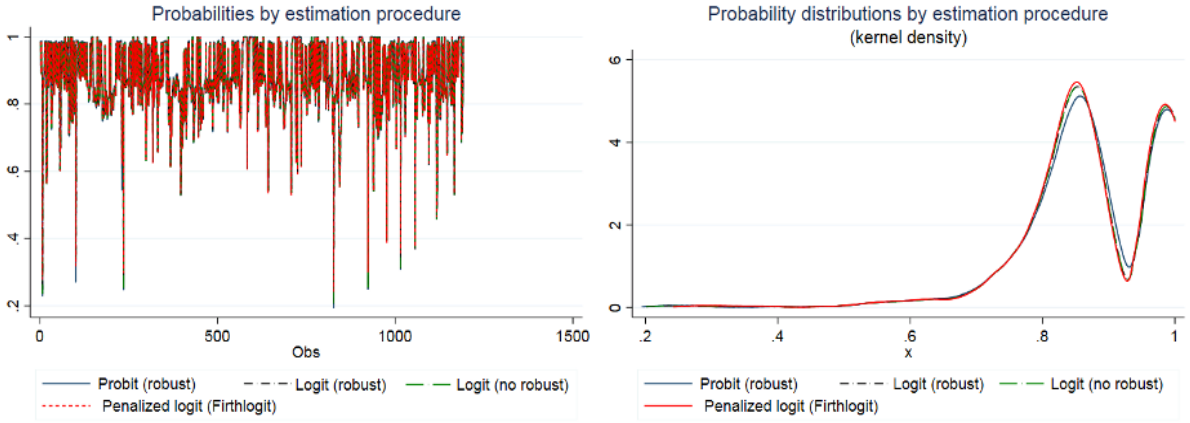

Figure E.3 — The different estimation procedures and the simulated probabilities

E.4. Variable correlations

Table E.4 — Correlations between the regressors

	Horizon/Premium	ER. Volatility	Reserves	Outlier	Round 2
Fixed ERR sample					
Horizon/Premium	1.0000				
ER. Volatility	-0.3943	1.0000			
	(0.000)				
Reserves	-0.0117	0.0157	1.0000		
	(0.642)	(0.533)			
Outlier	-0.1547	0.1340	-0.0612	1.0000	
	(0.000)	(0.000)	(0.015)		
Round 2	0.0684	-0.1132	-0.0844	-0.2733	1.0000
	(0.007)	(0.000)	(0.001)	(0.000)	
Lower Intermediate ERR sample					
Horizon/Premium	1.0000				
ER. Volatility	-0.3369	1.0000			
	(0.000)				
Outlier	-0.3418	0.2651	.	1.0000	
	(0.000)	(0.000)	.		
Round 2	0.3222	-0.2747	.	-0.3193	1.0000
	(0.000)	(0.0000)	.	(0.000)	
Upper Intermediate ERR sample					
Horizon/Premium	1.0000				
ER. Volatility	-0.6489	1.0000			
	(0.000)				
Reserves	0.0745	-0.1575	1.0000		
	(0.006)	(0.000)			
Outlier	-0.3304	0.6109	-0.1329	1.0000	
	(0.000)	(0.000)	(0.000)		
Round 2	0.2447	-0.3921	-0.3864	-0.0993	1.0000
	(0.000)	(0.000)	(0.000)	(0.000)	
(Full) Intermediate ERR sample					
Horizon/Premium	1.0000				
ER. Volatility	-0.3525	1.0000			
	(0.000)				
Reserves	-0.0017	-0.0744	1.0000		
	(0.942)	(0.001)			
Outlier	-0.2718	0.2927	-0.1749	1.0000	
	(0.000)	(0.000)	(0.000)		
Round 2	0.2215	-0.2408	-0.2494	-0.1660	1.0000
	(0.007)	(0.000)	(0.000)	(0.000)	
Floating ERR sample					
Horizon/Premium	1.0000				
ER. Volatility	-0.4993	1.0000			
	(0.000)				
Reserves	-0.1438	-0.3083	1.0000		
	(0.000)	(0.000)			
Outlier	-0.1630	0.4501	-0.2523	1.0000	
	(0.000)	(0.000)	(0.000)		
Round 2	0.1822	-0.3097	0.0272	-0.0574	1.0000
	(0.000)	(0.000)	(0.356)	(0.0477)	

Note: *p*.values are reported in parentheses.

E.5. Identification of the sample-specific sources of disagreements

E.5.1. Summary of the sequential approach results

Table E.5.1.1 — *Fixed* ERR sample

Dependent variable:		**Y = 0**		**Y = 1**	
No. observations:		979		590	
(Percentage)		(62.40)		(37.60)	
Ω		**TN**	**FP**	**TP**	**FN**
		$Pr^{adj}(\hat{Y}_i) = \Omega$			
Model:					
All factors	Obs.	781	198	390	200
Pseudo-R^2 = 0.1545	(%)	(79.78)	(20.22)	(66.10)	(33.90)
Excluding ...					
Horizon and/or Premium	Obs.	765	84	366	194
Pseudo-R^2 = 0.1381	(%)	(97.95)	(42.42)	(93.85)	(97.0)
E.R. volatility	Obs.	602	187	388	177
Pseudo-R^2 = 0.1290	(%)	(77.08)	(94.44)	(99.49)	(88.50)
Reserves	Obs.	781	196	354	199
Pseudo-R^2 = 0.0698	(%)	(100)	(98.99)	(90.77)	(99.50)
Round 2	Obs.	460	198	309	22
Pseudo-R^2 = 0.1298	(%)	(58.90)	(100)	(79.23)	(11.0)

Notes: The percentage of the observations in models excluding a variable are calculated relative to the number of observations in the full models —i.e. "All factors".

Table E.5.1.2 — *Lower-Intermediate* ERR sample

Dependent variable:		**Y = 0**		**Y = 1**	
No. observations:		460		434	
(Percentage)		(51.45)		(48.55)	
Ω		**TN**	**FP**	**TP**	**FN**
		$Pr^{adj}(\hat{Y}_i) = \Omega$			
Model:					
All factors	Obs.	410	50	241	193
Pseudo-R^2 = 0.1471	(%)	(89.13)	(10.87)	(55.53)	(44.47)
Excluding ...					
Horizon and/or Premium	Obs.	407	10	227	191
Pseudo-R^2 = 0.1183	(%)	(99.27)	(20.0)	(94.19)	(98.96)
Outlier	Obs.	410	43	225	193
Pseudo-R^2 = 0.1367	(%)	(100)	(86.0)	(93.36)	(100)
Round 2	Obs.	261	21	129	102
Pseudo-R^2 = 0.0888	(%)	(63.66)	(42.0)	(53.53)	(52.85)

Notes: The percentage of the observations in models excluding a variable are calculated relative to the number of observations in the full models —i.e. "All factors".

Table E.5.1.3 — *Upper-Intermediate* ERR sample

Dependent variable:		**Y = 0**		**Y = 1**	
No. observations:		460		891	
(Percentage)		(34.05)		(65.95)	
Ω		**TN**	**FP**	**TP**	**FN**
		$Pr^{adj}(\hat{Y}_i) = \Omega$			
Model:					
All factors	Obs.	450	10	833	58
Pseudo-R^2 = 0.8279	(%)	(97.83)	(2.17)	(93.49)	(6.51)
Excluding ...					
Horizon and/or Premium	Obs.	450	0	806	57
Pseudo-R^2 = 0.8058	(%)	(100.0)	(0.0)	(96.76)	(98.28)
E.R. volatility	Obs.	450	6	820	57
Pseudo-R^2 = 0.8103	(%)	(100.0)	(60.0)	(98.44)	(98.28)
Reserves	Obs.	444	5	561	50
Pseudo-R^2 = 0.2301	(%)	(98.67)	(50.0)	(67.35)	(86.21)
Outlier	Obs.	450	1	808	57
Pseudo-R^2 = 0.8218	(%)	(100.0)	(10.0)	(97.0)	(98.28)
Round 2	Obs.	449	1	807	57
Pseudo-R^2 = 0.7960	(%)	(99.78)	(10.0)	(96.88)	(98.28)

Notes: The percentage of the observations in models excluding a variable are calculated relative to the number of observations in the full models —i.e. "All factors".

Table E.5.1.4 — Full *Intermediate* ERR sample

Dependent variable:		**Y = 0**		**Y = 1**	
No. observations:		460		1481	
(Percentage)		(23.70)		(76.30)	
Ω		**TN**	**FP**	**TP**	**FN**
		$Pr^{adj}(\hat{Y}_i) = \Omega$			
Model:					
All factors	Obs.	430	30	1200	281
Pseudo-R^2 = 0.4482	(%)	(93.48)	(6.52)	(81.03)	(18.97)
Excluding ...					
Horizon and/or Premium	Obs.	361	29	1196	209
Pseudo-R^2 = 0.4172	(%)	(83.95)	(96.67)	(99.67)	(74.38)
Reserves	Obs.	415	28	907	251
Pseudo-R^2 = 0.1762	(%)	(96.51)	(93.33)	(75.58)	(89.32)
Outlier	Obs.	430	9	1026	280
Pseudo-R^2 = 0.4430	(%)	(100.0)	(30.0)	(85.50)	(99.64)
Round 2	Obs.	421	7	967	256
Pseudo-R^2 = 0.3466	(%)	(97.91)	(23.33)	(80.58)	(91.10)

Notes: The percentage of the observations in models excluding a variable are calculated relative to the number of observations in the full models —i.e. "All factors".

E.5.1.5 — *Floating* ERR sample

Dependent variable:		**Y = 0**		**Y = 1**	
No. observations:		146		1047	
(Percentage)		(12.24)		(87.76)	
Ω		**TN**	**FP**	**TP**	**FN**
		$Pr^{adj}(\hat{Y}_i) = \Omega$			
Model:					
All factors	Obs.	123	23	731	316
Pseudo-R^2 = 0.1546	(%)	(84.25)	(15.75)	(69.82)	(30.18)
Excluding ...					
Horizon and/or Premium	Obs.	109	13	604	225
Pseudo-R^2 = 0.1238	(%)	(88.62)	(56.52)	(82.63)	(71.20)
E.R. volatility	Obs.	105	10	546	245
Pseudo-R^2 = 0.1246	(%)	(85.37)	(43.48)	(74.69)	(77.53)
Reserves	Obs.	119	14	571	277
Pseudo-R^2 = 0.0733	(%)	(96.75)	(60.87)	(78.11)	(87.66)
Round 2	Obs.	120	19	685	293
Pseudo-R^2 = 0.1193	(%)	(97.56)	(82.61)	(93.71)	(92.72)

Notes: The percentage of the observations in models excluding a variable are calculated relative to the number of observations in the full models —i.e. "All factors".

Table E.5.1.6 — The sample-specific sources of disagreements (summary)

Variables	Estimation sample				
	Fixed	Intermediate			Float
		Lower	*Upper*	*Full*	
ER. Volatility	2 (0.34)	—	12 (1.35)	—	123 (11.75)
Horizon and/or Premium	22 (3.73)	14 (3.23)	15 (1.68)	—	90 (8.60)
Multiple	333 (56.44)	212 (48.85)	550 (61.73)	94 (60.26)	432 (41.26)
Outlier	—	15 (3.46)	1 (0.11)	3 (1.92)	—
Reserves	33 (5.59)	—	255 (28.62)	35 (22.44)	62 (5.92)
Round 2	—	—	—	—	24 (2.29)
Total model (TP)	390 (66.10)	241 (55.53)	833 (93.49)	133 (85.26)	731 (69.82)
Diff. in thresholds (FN)	200 (33.90)	193 (44.47)	58 (6.51)	23 (14.74)	316 (30.18)
Total (TP+FN)	590	434	891	156	1047

Note: Entries correspond to the frequencies of the occurrence. Percentages (of the total number of occurrence) are reported in parentheses. Omitted variables have 0 occurrence —or have been discarded following the likelihood ratio test. FN (resp. TP) stands for false negative (resp. true positive).

E.5.2. Auxiliary analyses

Table E.5.2.1 — LR tests

Excluded variable	Fixed	Intermediate			Float
		Lower	*Upper*	*Full*	
Horizon and/or Premium	33.94 [0.000]	35.70 [0.000]	38.36 [0.000]	58.83 [0.000]	27.32 [0.000]
E.R. volatility	52.94 [0.000]	3.29 [0.069]	30.56 [0.000]	0.65 [0.421]	26.64 [0.000]
Reserves	175.97 [0.000]		1035.95 [0.000]	574.09 [0.000]	72.11 [0.000]
Outlier	2.59 [0.1074]	12.86 [0.001]	10.65 [0.001]	11.81 [0.001]	0.00 [0.999]
Round 2	51.25 [0.000]	146.55 [0.000]	55.35 [0.000]	206.12 [0.000]	31.36 [0.000]

Null: $Model_X$ is nested in $Model_{All_factors}$. p.values are reported in brackets.

Table E.5.2.2 — Models and ROC area difference tests (*Fixed* ERR sample)

Model	ROC Area	Std. Err.	χ^2	df	$Pr. > \chi^2$	Bonferroni $Pr. > \chi^2$
All factors (standard)	0.7579	0.0124				
Horizon and/or Premium	0.8113	0.0111	52.7827	1	0.0000	0.0000
ER. volatility	0.6965	0.0138	109.4397	1	0.0000	0.0000
Reserves	0.7227	0.0131	32.9347	1	0.0000	0.0000
Round 2	0.7697	0.0115	1.8496	1	0.1738	0.6953

Note: "Std. Err." stands for standard error. χ^2 indicates the chi-squared statistics associated to the test. "*df*" stands for degree of freedom. The different areas are compared to the full model ROC area.

Table E.5.2.3 — Models and ROC area difference tests (*Lower -Intermediate* ERR sample)

Model	ROC Area	Std. Err.	χ^2	df	$Pr. > \chi^2$	Bonferroni $Pr. > \chi^2$
All factors (standard)	0.7251	0.0176				
Horizon and/or Premium	0.8114	0.0144	53.0203	1	0.0000	0.0000
Outlier	0.7218	0.0176	0.8809	1	0.3480	1.0000
Round 2	0.5620	0.0193	59.4244	1	0.0000	0.0000

Note: "Std. Err." stands for standard error. χ^2 indicates the chi-squared statistics associated to the test. "*df*" stands for degree of freedom. The different areas are compared to the full model ROC area.

Table E.5.2.4 — Models and ROC area difference tests (*Upper-Intermediate* ERR sample)

Model	ROC Area	Std. Err.	χ^2	*df*	$Pr. > \chi^2$	Bonferroni $Pr. > \chi^2$
All factors (standard)	0.9913	0.0017				
Horizon and/or Premium	0.9856	0.0026	12.8243	1	0.0003	0.0017
ER. volatility	0.9885	0.0021	4.6350	1	0.0313	0.1566
Reserves	0.7824	0.0125	286.5396	1	0.0000	0.0000
Outlier	0.9901	0.0018	5.9568	1	0.0147	0.0733
Round 2	0.9898	0.0018	4.1107	1	0.0426	0.2131

Note: "Std. Err." stands for standard error. χ^2 indicates the chi-squared statistics associated to the test. "*df*" stands for degree of freedom. The different areas are compared to the full model ROC area.

Table E.5.2.5 — Models and ROC area difference tests (Full *Intermediate* ERR sample)

Model	ROC Area	Std. Err.	χ^2	*df*	$Pr. > \chi^2$	Bonferroni $Pr. > \chi^2$
All factors (standard)	0.9085	0.0065				
Horizon and/or Premium	0.8870	0.0066	30.6812	1	0.0000	0.0000
Reserves	0.7732	0.0110	264.9763	1	0.0000	0.0000
Outlier	0.9069	0.0066	1.9547	1	0.1621	0.6483
Round 2	0.8312	0.0088	90.7793	1	0.0000	0.0000

Note: "Std. Err." stands for standard error. χ^2 indicates the chi-squared statistics associated to the test. "*df*" stands for degree of freedom. The different areas are compared to the full model ROC area.

Table E.5.2.6 — Models and ROC area difference tests (*Floating* ERR sample)

Model	ROC Area	Std. Err.	χ^2	*df*	$Pr. > \chi^2$	Bonferroni $Pr. > \chi^2$
All factors (standard)	0.8160	0.0151				
Horizon and/or Premium	0.7939	0.0173	2.7442	1	0.0976	0.3904
ER. volatility	0.7382	0.0175	32.1552	1	0.0000	0.0000
Reserves	0.7296	0.0176	84.8478	1	0.0000	0.0000
Round 2	0.8194	0.0146	0.2420	1	0.6228	1.0000

Note: "Std. Err." stands for standard error. χ^2 indicates the chi-squared statistics associated to the test. "*df*" stands for degree of freedom. The different areas are compared to the full model ROC area.

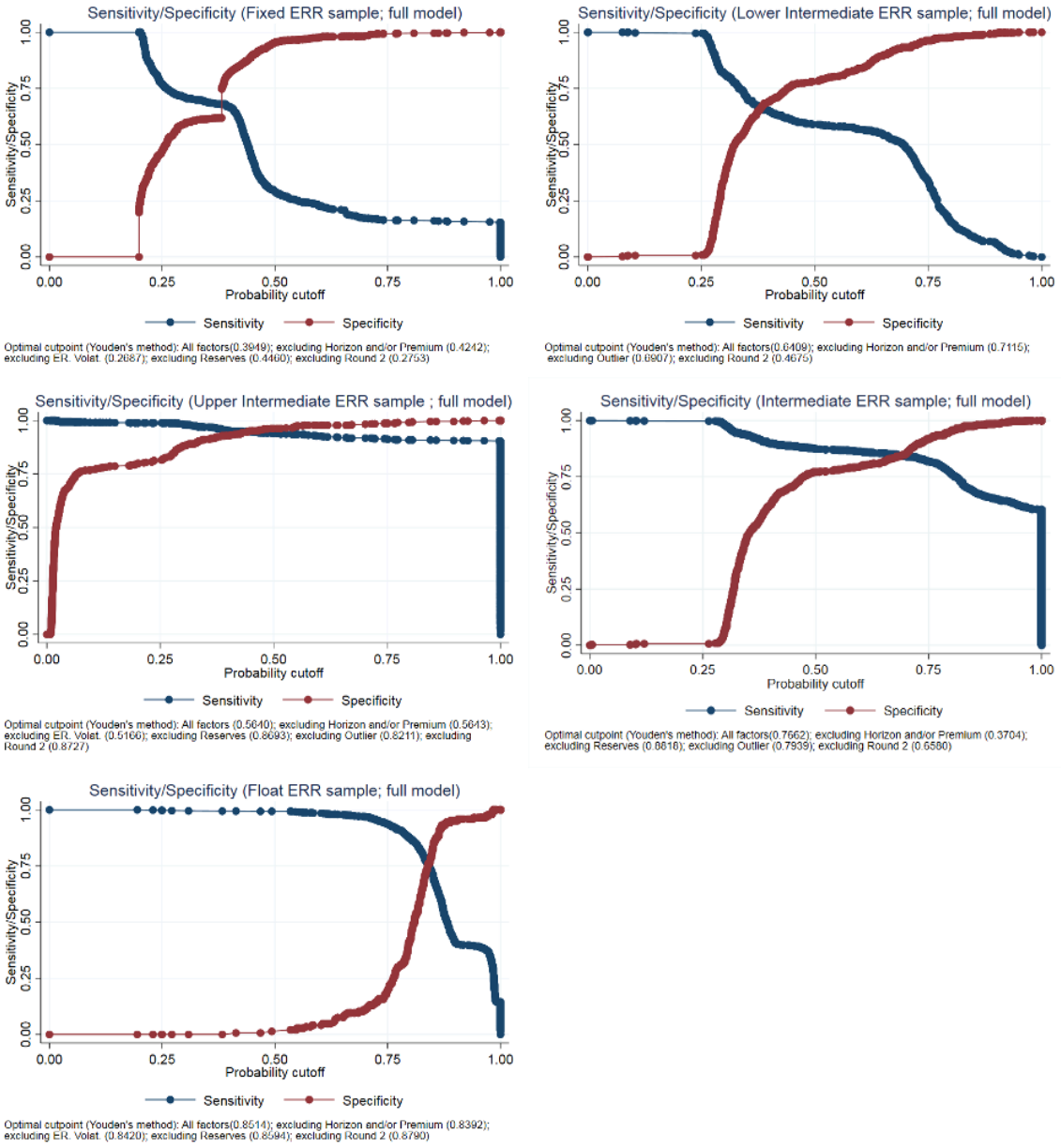

Figure E.5.2.1 — Sensitivity/Specificity *vs.* probability cutoff (by estimation sample)

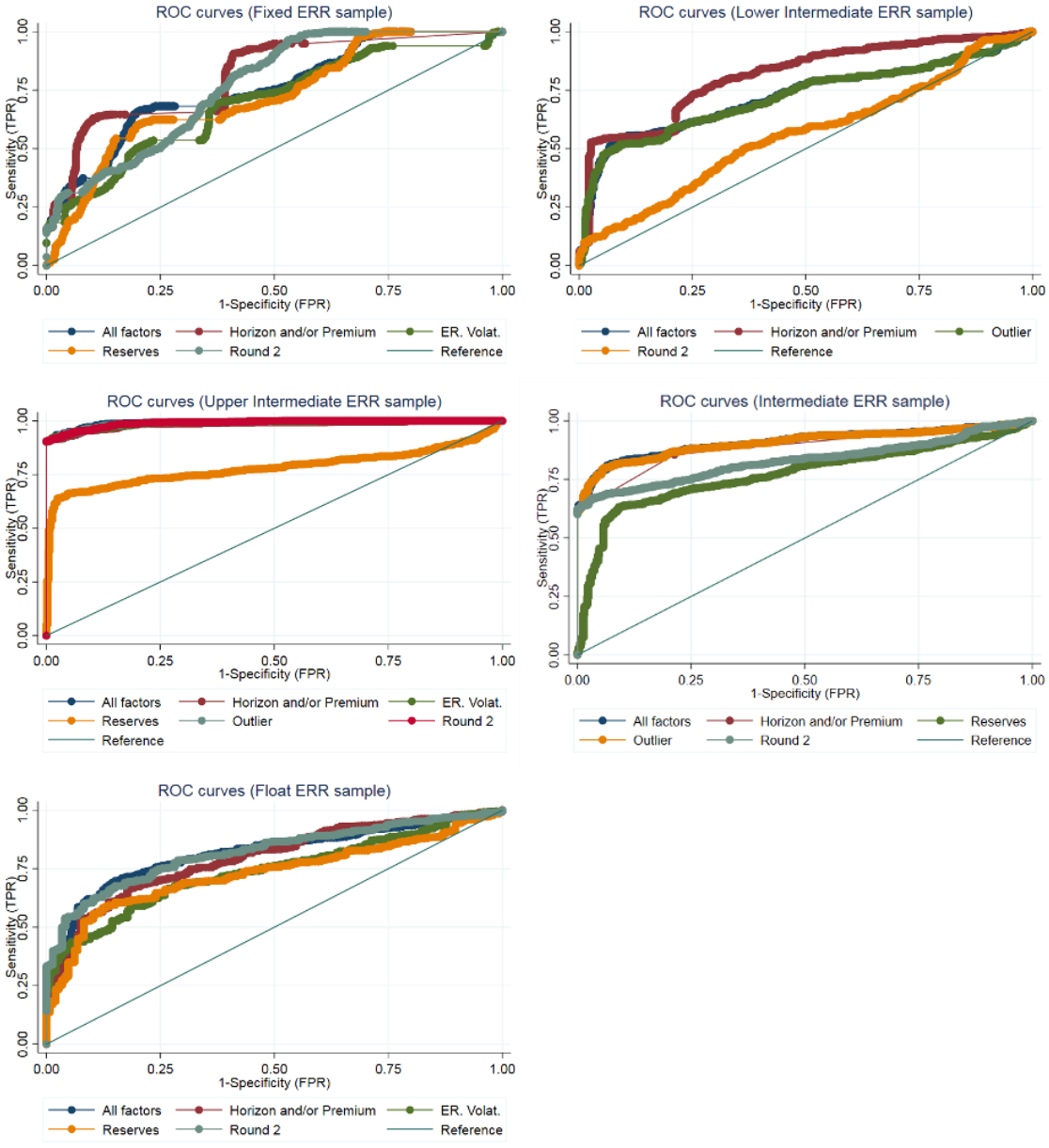

Figure E.5.2.2 — Models and ROC areas (by estimation sample)