MAFE Final Conference

Measuring Information Asymmetries and Modeling their Impact on Senegalese Migrants' Remittances

Marlon SEROR*

Paris School of Economics

Abstract

The literature on migrants' remittances commonly bypasses information asymmetry issues, arguing that transfers occur within the extended family, and thus altruism should be strong enough to avoid the pitfalls that foreign investors encounter. We use novel, matched data on Senegalese migrants (in France, Italy and Mauritania) and their households of origin to measure empirically discrepancies in migrants' and households' responses to identical survey questions. As the discrepancies prove to be systematically non–zero, we assess potential explanations and provide evidence that information asymmetry is most likely at work. Testable predictions are obtained from a simple theoretical framework and supported by the data. A natural question is then whether information asymmetry affects empirical analyses of transfer determinants. We derive the bias induced by information asymmetry in transfer regressions thanks to the model, and show that information asymmetry may entail a serious bias, which is difficult to sign *a priori*.

*E–mail: marlon.seror@ens.fr

Acknowledgements

This paper builds upon the Master's Thesis I wrote at the Paris School of Economics and the University of California, Berkeley. I would thus like to thank my supervisors, Flore Gubert and Ted Miguel. I am also indebted to Jenny Aker, Marshall Burke, Manon Falquerho, Émilie Hu, Hu Peigen, Guillaume Roger, Delphine Roy, Jean–Noël Senne and Miri Stryjan.

Contents

1 Introduction

1.1 Research question

Migrants' remittances are now "three times the size of official development assistance" (World Bank, 2011). This evolution has aroused enthusiasm, partly because migrants are expected to have developed family networks and home town associations allowing them to process information from their origin countries more effectively than foreign institutions or investors (Osili, 2004). Put differently, migrants should be able to overcome the informational problems that make foreign endeavors risky as "altruism facilitates the reinforcement of arrangements that are mutually beneficial" (Carling, 2008).

This presupposition was not challenged until recently. The economic literature recognizes that "various types of informational asymmetries may arise in the context of migration" (Rapoport and Docquier, 2006) because "once the migrant is abroad, an informational asymmetry is created in favor of the non–migrants with respect to the economic conditions at home" (*ibid.*). But little work has been done in this direction, mostly because of data constraints, and information asymmetries are either disregarded, papers treating migrants' and households' responses to surveys as equivalent, or when their impact on remittances is under study, they are *assumed* rather than *shown*.[1]

Such a "gap" is pregnant with theoretical implications for the unitary and collective models of intra–household decision–making: "A leading candidate explanation for observed inefficiencies is asymmetry of information in the household, so that family members cannot monitor each other well enough to enforce mutually–beneficial cooperative agreements" (Ashraf et al., 2011).

Besides, there might be policy implications to a divergence between migrant and household. On the one hand, if information asymmetry is construed as rent extraction, the more opaque the use of the remittances back home, the more households of origin may be able to extract; on the other, if migrants realize that and can afford to remit less,[2]

[1] Azam and Gubert (2004) defend the idea that insurance contracts through remittances might be fraught with moral hazard issues, even within a family, in the following way: "One could argue that family members share naturally a great deal of information, so that the informational asymmetry assumption is untenable. However, most migrants in our sample live in France and only visit their home village every third year on average. This leaves a lot of room for asymmetric information."

[2] *I.e* if they can bear the consequences of an underfinanced household of origin. This is all the more

serious information asymmetry likely leads to inefficiently low transfers and fewer (or poorly implemented) investment projects by the migrant at the extensive (intensive) margin. How this translates into policy recommendations depends on whose projects are more "valuable"—the remittance sender's or the recipient's.

1.2 Literature review

Recent studies suggest that migrants' transfers indeed occur in an imperfect–information world. Dia (2007) shows on the basis of ethnographic surveys that "conflicts with relatives over money and its use" may emerge as the household in Senegal "had rather keep the money [...], would not purchase all that is necessary, or would use up the transfers in a short time and thus imperil the household's food security".[3]

However, most existing evidence in the economic literature is based on reduced–form empirical strategies or predictions of information–asymmetry models that are supported by the data, but need not be inconsistent with other explanations.

Azam and Gubert (2004) show empirically that moral hazard may arise from co–insurance contracts between migrants and non–migrants, as the well–known weakening of incentives that derives from insurance with imperfectly observable effort is supported by their data from the Kayes region of Mali. They find indeed evidence of "shirking" for farmers insured by one or several migrants.[4] This suggests that "geographical dispersion [...] makes it difficult to monitor performance and creates moral hazard" so that "each participant in the insurance pool has the incentive to underreport income or to reduce effort" (*ibid.*). But however plausibly moral hazard may explain the lower effort observed in households with emigrants, it is not directly proved.

Focusing not on insurance contracts but on transfer and monitoring behaviors among Kenyan "split couples" (husbands living temporarily in Nairobi, while their wives and children stay in the village of origin), De Laat (2008) intends to test "for the endo-

relevant as the primary remittance motive is "exchanges of various services" in the face of "pervasive credit market imperfections" (Rapoport and Docquier, 2006).

[3] Translation by the author. In order to avoid such conflicts, "migrants allocate money transfers beforehand. Instead of sending the money directly to the household, they send it to shopkeepers." Dia (2007) also mentions an interviewee who planned to open a drugstore in Senegal with a cousin of his. The cousin embezzled the monies, and the interviewee found out about it as "he discreetly investigated" during vacations in Senegal.

[4] *I.e.* "the more reliable is the insurance mechanism, after adjusting for risk, the higher is the incentive to shirk, and the lower is the expected output" (Azam and Gubert, 2004).

geneity of one of the main elements separating the different models of household allocations, namely the household information set". Indeed, the assumption of full and symmetric information is innocuous if it can be argued a "by–product of daily interactions" (*ibid.*), but these are precisely what migration prevents. De Laat cannot come to a definitive conclusion on the "impact of increased monitoring on remittances". Even if we believe in the identification strategy,[5] he bypasses information asymmetry and can thus only investigate the relationship between monitoring and transfers, *i.e.* a reduced form. Now, if monitoring means increased communication (whether by phone or through visits to the rural wife), the empirical strategy needs to check whether the migrant does not call or meet *additional* relatives, whom he decides to send money to, care of his "older" recipients. Another potential confound is that improved communication enables the migrant to better perceive the recipients' needs and might thus as a consequence increase transfers. Nevertheless, an interesting contribution of De Laat (2008) for our research is the considerable amounts spent by Kenyan urban migrants to visit, communicate and, arguably, monitor their rural wives,[6] so that significant efficiency losses arise as "increases in the cost of monitoring reduce the benefits of split migration relative to no– and joint migration" (*ibid.*).

Efficiency would then be improved if the same decisions were arrived at while sparing the migrant the cost of a constant watch. Ashraf et al. (2011) find using a randomized control trial among the Salvadoran community in Washington, D.C., that migrants are more likely to open savings accounts in El Salvador into which they could send remittances if they offer greater control. This suggests that migrants feel the need to exert more control over remittances, which in turn hints at non–trivial information asymmetries.[7] Yet the empirical analysis is again a "reduced–form": What

[5] De Laat (2008) uses proxies for monitoring costs as instruments for monitoring variables in transfer regressions. However, the exclusion restriction might not be respected. In effect, "Transport time to home/spouse" is a function of the migration decision, and one of the main tenets of the "new economics of labor migration" since Lucas and Stark (1985) is that "decisions about remittances are linked with decisions about migration" (Carling, 2008). Now, if rural husbands who live very far from Nairobi only engage in labor migration with the project of remitting most of their income to compensate for the cost of settling far from home, the exclusion restriction may not be respected.

[6] A limitation of De Laat's survey seems to be a failure to disentangle monitoring from simple altruism (the migrant misses his wife and children), a criticism that we shall bear in mind for Sections 3 and 5.

[7] Indeed, in their baseline survey 20.7% of the interviewees claimed being "interested in direct payments to improve control over remittance uses" (Ashraf et al., 2011).

Ashraf et al. (2011) measure is the effect of monitoring costs on the demand for savings/remittance accounts, not the presence of information asymmetry or demand for monitoring—however probable these interpretations. By their own admission, the results are consistent "with migrants exerting control over savings in part to build up buffer stocks (precautionary savings) that need to be accessed quickly by primary remittance recipients in case of emergency" (*ibid.*). In other words, the observed increase in take–up as control is improved could be due to the lack of secure saving mechanisms available to the migrants, not to a lack of trust in their households of origin.

Finally, Chort et al. (2012) provide suggestive evidence that migrant associations are used by households of origin to monitor their emigrant members, maybe by manipulating reputations and spreading rumors (remittances are in their model a fee to access network services). But one cannot exclude that migrants themselves rely on them for information. Membership in associations that bring together migrants from the same village or neighborhood of origin could therefore be a further hint that information asymmetries between remittance senders and recipients exist and are pervasive.

The issue with this emerging literature is that, even when it makes use of an experimental identification strategy, it is still a "reduced form":[8] The particular channel through which costlier or improved "monitoring" affects transfers need not be information asymmetry. In order to show that this particular channel is at work, one needs to directly measure information asymmetries. For this we need "matched" data, *i.e.* migrants' and origin households' responses to comparable (and quantifiable) survey questions. But due to sizeable collection costs, survey data are usually implemented either in the host or the origin country of migrants, so misrepresentations cannot be captured. For instance, De Laat's "survey did not interview rural women".[9]

In the absence of information asymmetry, "one–sided" data may only induce mea-

[8]Transfers are regressed on monitoring variables instrumented by proxies for monitoring costs (De Laat, 2008), or on treatments that vary exogenously the degree of monitoring that the migrant can exert (Ashraf et al., 2011).

[9]The Mexican Migration Project does collect data both in home (Mexico) and host (U.S.) countries, but at community–, not family–level: Households are randomly surveyed in Mexican communities; in a second stage, migrants from the same communities are interviewed in destination areas, on a nonrandom basis; they are not necessarily relatives of the Mexico interviewees, as the purpose of the U.S. survey is rather to reach "migrants who have established their households in the United States". See `http://mmp.opr.princeton.edu/databases/studydesign-en.aspx`.

surement error. In this perspective, Osili (2004, 2007) collected matched data in order to combine in one transfer equation migrants' and households' reports of their own characteristics.[10] The "US–Nigeria Migration Study" consisted of sampling 120 Igbo Nigerian migrants registered in the Chicago phonebook, surveying them and then (if the migrants agreed to provide their contacts) their households of origin in Nigeria. The final sample comprised 61 migrant and origin–family pairs. But unfortunately, Osili (2004, 2007) does not contrast migrants' and households' reports on similar survey items. Now, if information asymmetries *are* an issue, there is no obvious reason to prefer nonmigrants' over migrants' reports of the former's characteristics, and if diverging reports are likely to be a by–product of rent extraction, then the transfer regressions commonly run in the literature might be vitiated by reverse causation—as migrants' reports would reflect the household's misrepresentations and extraction of extra remittances.

The next section presents the data and summary statistics. Section 3 investigates the discrepancies between migrants' and households' reports and assesses plausible explanations. Section 4 develops a soft–information model and isolates a few predictions tested in Section 5 to further buttress the interpretation of the discrepancies as evidence of rent extraction. The model is then used in Section 6 to derive the potential bias in transfer regressions from ignoring information asymmetry. Finally, Section 7 provides a few recommendations for the study of migrants' remittances. Section 8 concludes.

2 "Matched" data

2.1 Presentation of the data

In order to investigate information asymmetry in migrant remittances, we can draw on the "Migration and development in Senegal: an empirical analysis using matched data on Senegalese migrants and their origin households" (henceforth, MIDDAS) project (2009–2010).[11] As the data in Osili (2004, 2007), MIDDAS is based on a sample of migrants contacted in the host country. But the survey design aspired to a nation-

[10] She also claims that matched data can help us partially overcome the problematic impact of "migrant's past transfers on the home household's asset holdings" (Osili, 2004).

[11] For a detailed presentation of the project, see `http://www.dial.ird.fr/projets-de-recherche/projets-anr/middas` [in French].

ally representative sample of the Senegalese immigrant population in France, Italy and Mauritania.[12] Although the pitfall of a matched sample smaller than the initial one cannot be avoided, the initial MIDDAS samples are of respectable sizes: 300 surveys carried out for the French sample of migrants, 302 in Italy and 326 in Mauritania. As far as the matched sample is concerned, 95 households could be interviewed corresponding to the French sample and 62 to the Italian one. Thanks to geographical proximity notably, 174 households of origin of Senegalese migrants in Mauritania could be surveyed. The households interviewed in Senegal were presented with the "Pauvreté et Structure Familiale" (PSF) survey (see De Vreyer et al. (2008) for a description of the data), a thorough questionnaire that was first implemented on a nationally representative sample of the Senegalese population. A feature of "migration–PSF" (the questionnaire is the same as the original PSF, but the population differs) is its painstaking description of the complex structure of Senegalese households, made up of several nuclear groups.

Thanks to the matched structure of the data, it is possible to compare migrants' and households' reports (X_{hm} and X_{hh}, respectively) of a set of household characteristics X_h , and thus *measure* gaps in m's and h's information sets.[13] Unfortunately, the structure of the data does not enable us to test for information asymmetries created by the migrant.

Because the focus of MIDDAS was not only information asymmetries, a short questionnaire was created in order to delve further into this issue, and implemented in October 2012. The questionnaire is made of three parts. The first section follows MIDDAS in briefly asking the migrant about her remittances.[14] The second part is meant to assess whether information asymmetry is a concern for the migrant.[15] Last but not

[12] The same survey was implemented in Côte d'Ivoire but it will not be used here because no matching was carried out with those data. For an overview of the sampling procedure, please refer to Chort et al. (2012).

[13] To my best knowledge, empirical measures of asymmetric information are only to be found in the finance literature. Brown et al. (2004), for instance, use the "ex ante probability that the first trade of the day is based on private information". Contrary to what MIDDAS enables us to do, this is no direct comparison of information among market makers. It is based on straightforwardly interpretable actions (buying or selling), which are not easily transposed into the remittance framework.

[14] *I.e.* whether she sends remittances to Senegal, to which household she sends most, how much, how often, and what their relationship is.

[15] It contains detailed questions on earmarking of transfers for particular purchases or people within the household, on the migrant's suspicion of misrepresentations, on monitoring by the migrant of the household's actions and situation, and communication more generally, on the possibility of sanction in case misrepresen-

least, econometric results from the comparison of MIDDAS and migration–PSF data (presented in Section 3) were submitted to the interviewees, who were asked to put forward what in their views are the most plausible hypotheses.

Great care was taken in letting the interviewees express themselves and in gathering as much qualitative information as possible.[16] Due to the qualitative dimension of the survey, sample size consists of 20 observations. Besides, the population studied is different than that in MIDDAS because of limited time and financial resources: Respondents were randomly approached in the streets of the Château Rouge area in Paris, which is densely populated by Senegalese migrants and attracts many others from the rest of the Paris region.

2.2 Descriptive statistics

Selection into the matched sample is not crucial to our argument. If anything, the matched sample should overrepresent migrant–household pairs with a good relationship and thus constitute a lower bound in terms of information asymmetry.[17] However, we shall provide a number of empirical tests, as selection could harm external validity.

First, Tables 1 through 3 present descriptive statistics and investigate selection into the matched samples for migrants living in France, Italy and Mauritania, respectively. The only significant (at the 90% level) divergence between the French unmatched and matched samples is household size, slightly higher in the latter. The Italian matched sample appears more selected: They are less often Wolof,[18] come from larger, wealthier households, and display significantly fewer missing values for transfer amount variables. Conversely, the Mauritanian matched sample is more often Wolof, but they also come from larger (and wealthier, although by a different measure) households, and are less likely not to report amounts remitted. They tend to have fewer recipients and

tations are detected, and on whether the migrant too may distort the information she provides her relatives in Senegal.

[16] For instance, anecdotes told by migrants, or whether they answered spontaneously rather than waiting for the enumerator to provide a list of options, were conscientiously recorded.

[17] In an earlier work (Seror, 2012), we provided suggestive evidence that misrepresentations by h are differentially more of a problem in the unmatched than in the matched sample, vindicating the lower–bound hypothesis. Note that Seror (2012) does not take heteroskedasticity into account. But few results change when robust standard errors are used. Heteroskedasticity–free regressions are available upon request.

[18] The significance of the "Wolof" indicator variable probably captures the ethnic diversity of the Senegal River basin, as opposed to the predominantly Wolof immigrant population in Italy. Now, the Senegal River basin is mostly rural and therefore households were easier to find than in big cities.

marginally more of them have a child staying with their relatives in Senegal. Finally, there are more remittance senders in the matched sample.

The descriptive statistics tables suggest that selection into the matched sample is not a serious issue. This is corroborated—for each country and over the whole sample—by Table 4, which displays Probit regressions of an indicator variable equal to 1 if the migrant's family could be successfully surveyed on migrant and household, as well as enumerator, characteristics. The rationale for including the latter set of variables is that the migrant's decision to provide a (truthful) contact for his or her household of origin may have been a function of the surveyor's observable characteristics.[19]

For France (Column 1), the migrant's age and total transfers to Senegal are significantly higher, and monetary transfers lower, in the matched sample. The Italian matched sample (Col. 2) is slightly more selected. None of the differences in the Mauritanian sample (Col. 3) are significant at more than the 90% level. In the Italian sample, enumerator gender seems to have played a marginal role: The negative coefficient suggests that (predominantly male) migrants were less likely to put female enumerators in touch with their households of origin. Finally, when the same exercise is done for the whole sample (Col. 4), we see that the coefficient on the Mauritanian dummy is positive and significant. This reflects the fact that "tracking" households of origin was easier from Mauritania due to geographical proximity.

To conclude on Tables 1 through 4, there is indeed some selection into the matched sample. However, it does not seem very serious.

[19] Enumerator characteristics and how they might have affected interviewees' responses is discussed in Section 3.2.

Table 1: Differences between matched and unmatched samples from France

	Unmatched sample	Matched sample	Difference	p-value
Migrant characteristics:				
Male (d)	0.75	0.78	-0.03	0.532
Migrant's age	36.60	37.84	-1.24	0.369
Size of migrant's household (in host country)	2.62	2.37	0.25	0.315
Ethnicity: Wolof (d)	0.30	0.22	0.08	0.171
Religion: Mourid (d)	0.20	0.23	-0.03	0.498
Religion: Tidjan (d)	0.25	0.22	0.03	0.590
Migration history:				
Year of arrival in host country	1994.37	1995.91	-1.55	0.248
Time since arrival in host country	14.65	13.09	1.56	0.244
Economic situation:				
m's per capita total household income (per month, €)	951.88	977.69	-25.81	0.772
Unemployed (d)	0.13	0.16	-0.03	0.485
Family situation:				
Number of children	1.83	1.66	0.17	0.530
Number of resident children	0.70	0.57	0.13	0.465
Spouse resident in *h* (d)	0.16	0.12	0.04	0.385
At least one child lives with *h* (d)	0.21	0.23	-0.03	0.606
Characteristics of the household of origin:				
Size of the household of origin	14.33	16.88	-2.55*	0.077
Household is richer than community (d)	0.09	0.09	0.00	0.958
Household is poorer than community (d)	0.19	0.25	-0.06	0.261
h's wealth score, *m*'s report	0.06	0.13	-0.07	0.787
Son/Daughter of *h*'s head (d)	0.62	0.68	-0.06	0.319
Remittance behavior:				
Remit in cash/kind to the household of origin (d)	0.81	0.87	-0.06	0.228
Number of recipient members in the household of origin	1.29	1.33	-0.04	0.755
Remit in kind to the household of origin (d)	0.28	0.25	0.02	0.669
Total transfers to *h* in the past 12 months (in kind incl., €)	2119.97	2509.90	-389.94	0.177
Total money transfers to *h* in the past 12 months, €	2072.01	2370.37	-298.36	0.285
Nb of coresidents of *m* who remit to *h*	1.46	1.22	0.24	0.528
Missing total transfers, in kind incl. (d)	0.09	0.07	0.02	0.555
Missing total money transfers (d)	0.09	0.05	0.03	0.352
Nb of non–coresidents who remit to *h*	2.19	2.05	0.13	0.702

The wealth is the first principal component of a vector of assets. (d) means the variable is a dummy. Remittance amounts are computed on the subsample of migrants with non–zero transfers. *m* (*h*) refers to the migrant (household of origin).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Differences between matched and unmatched samples from Italy

	Unmatched sample	Matched sample	Difference	p-value
Migrant characteristics:				
Male (d)	0.76	0.80	-0.04	0.510
Migrant's age	35.69	36.82	-1.13	0.351
Size of migrant's household (in host country)	3.17	3.25	-0.07	0.749
Ethnicity: Wolof (d)	0.81	0.66	0.15***	0.010
Religion: Mourid (d)	0.68	0.57	0.11	0.103
Religion: Tidjan (d)	0.25	0.25	0.00	0.961
Migration history:				
Year of arrival in host country	1999.20	1998.95	0.24	0.773
Time since arrival in host country	9.82	10.05	-0.23	0.783
Economic situation:				
m's per capita total household income (per month, €)	574.79	665.44	-90.66	0.173
Unemployed (d)	0.22	0.20	0.02	0.746
Family situation:				
Number of children	1.80	1.97	-0.16	0.617
Number of resident children	0.53	0.49	0.04	0.789
Spouse resident in *h* (d)	0.30	0.30	-0.00	0.948
At least one child lives with *h* (d)	0.36	0.39	-0.03	0.657
Characteristics of the household of origin:				
Size of the household of origin	10.12	12.68	-2.55**	0.016
Household is richer than community (d)	0.12	0.17	-0.05	0.335
Household is poorer than community (d)	0.11	0.15	-0.04	0.353
h's wealth score, *m*'s report	0.51	2.03	-1.52***	0.000
Son/Daughter of *h*'s head (d)	0.69	0.67	0.02	0.807
Remittance behavior:				
Remit in cash/kind to the household of origin (d)	0.74	0.80	-0.06	0.305
Number of recipient members in the household of origin	1.21	1.41	-0.20	0.239
Remit in kind to the household of origin (d)	0.19	0.25	-0.05	0.365
Total transfers to *h* in the past 12 months (in kind incl., €)	2269.57	2743.45	-473.88	0.137
Total money transfers to *h* in the past 12 months, €	2192.26	2662.60	-470.33	0.116
Nb of coresidents of *m* who remit to *h*	1.31	1.31	-0.00	0.994
Missing total transfers, in kind incl. (d)	0.20	0.05	0.15***	0.004
Missing total money transfers (d)	0.20	0.05	0.15***	0.004
Nb of non–coresidents who remit to *h*	1.63	2.05	-0.42	0.163

The wealth is the first principal component of a vector of assets. (d) means the variable is a dummy. Remittance amounts are computed on the subsample of migrants with non–zero transfers. *m* (*h*) refers to the migrant (household of origin).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Differences between matched and unmatched samples from Mauritania

	Unmatched sample	Matched sample	Difference	p-value
Migrant characteristics:				
Male (d)	0.61	0.66	-0.05	0.340
Migrant's age	34.91	36.69	-1.78	0.136
Size of migrant's household (in host country)	3.14	3.05	0.09	0.728
Ethnicity: Wolof (d)	0.51	0.66	-0.15***	0.004
Religion: Mourid (d)	0.27	0.29	-0.02	0.650
Religion: Tidjan (d)	0.42	0.44	-0.02	0.748
Migration history:				
Year of arrival in host country	1989.05	2002.80	-13.75	0.266
Time since arrival in host country	6.98	6.20	0.78	0.371
Economic situation:				
m's per capita total household income (per month, €)	80.22	96.55	-16.34	0.208
Unemployed (d)	0.08	0.13	-0.05	0.152
Family situation:				
Number of children	2.26	2.55	-0.29	0.360
Number of resident children	1.11	0.92	0.19	0.327
Spouse resident in *h* (d)	0.17	0.25	-0.08	0.101
At least one child lives with *h* (d)	0.35	0.45	-0.10*	0.071
Characteristics of the household of origin:				
Size of the household of origin	9.37	10.57	-1.20*	0.085
Household is richer than community (d)	0.28	0.24	0.04	0.393
Household is poorer than community (d)	0.08	0.03	0.05**	0.042
h's wealth score, *m*'s report	-0.96	-0.75	-0.21	0.167
Son/Daughter of *h*'s head (d)	0.52	0.59	-0.06	0.259
Remittance behavior:				
Remit in cash/kind to the household of origin (d)	0.66	0.77	-0.11**	0.030
Number of recipient members in the household of origin	1.29	1.14	0.14**	0.031
Remit in kind to the household of origin (d)	0.31	0.31	0.00	0.975
Total transfers to *h* in the past 12 months (in kind incl., €)	593.36	745.58	-152.22	0.224
Total money transfers to *h* in the past 12 months, €	615.49	705.96	-90.47	0.493
Nb of coresidents of *m* who remit to *h*	1.79	1.23	0.57	0.156
Missing total transfers, in kind incl. (d)	0.10	0.06	0.05	0.117
Missing total money transfers (d)	0.11	0.05	0.06**	0.048
Nb of non–coresidents who remit to *h*	1.56	2.07	-0.51	0.150

The wealth is the first principal component of a vector of assets. (d) means the variable is a dummy. Remittance amounts are computed on the subsample of migrants with non–zero transfers. *m* (*h*) refers to the migrant (household of origin).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Determinants of selection into the matched sample

	(1) France	(2) Italy	(3) Mauritania	(4) All
Migrant characteristics:				
Migrant's age	0.0287**	0.0297	0.0142	0.0230***
	(0.0103)	(0.00922)	(0.00535)	(0.00891)
Son/Daughter of *h*'s head (d)	0.120	-0.0421	0.429*	0.197
	(0.0429)	(-0.0132)	(0.162)	(0.0758)
Migration history:				
Time since arrival in host country	-0.0166	-0.0167	0.00975	-0.00984
	(-0.00600)	(-0.00520)	(0.00367)	(-0.00382)
Economic situation:				
m's total monthly income (incl. social benefits, €)	-0.205	0.261**	0.228*	0.0992
	(-0.0740)	(0.0811)	(0.0859)	(0.0385)
Unemployed (d)	-0.272	0.681*	0.377	0.213
	(-0.0929)	(0.237)	(0.132)	(0.0839)
At least one child lives with *h* (d)	0.111	-0.0795	-0.0167	0.0292
	(0.0405)	(-0.0246)	(-0.00629)	(0.0113)
Spouse resident in *h* (d)	-0.451	0.0639	0.198	-0.0859
	(-0.150)	(0.0200)	(0.0733)	(-0.0331)
Characteristics of the household of origin:				
h's wealth score, *m*'s report	-0.00322	0.370***	0.0786	0.137**
	(-0.00116)	(0.115)	(0.0296)	(0.0532)
Household is poorer than community, *m*'s report (d)	0.210	0.0930	-0.647	-0.0239
	(0.0775)	(0.0295)	(-0.253)	(-0.00925)

Table 4: Determinants of selection into the matched sample (continued)

	(1) France	(2) Italy	(3) Mauritania	(4) All
Remittance behavior:				
Total transfers to *h* in the past 12 months (in kind incl., €)	1.147**	-0.124	1.009	0.412
	(0.413)	(-0.0386)	(0.380)	(0.160)
Total money transfers to *h* in the past 12 months, €	-0.962*	0.182	-1.052	-0.349
	(-0.347)	(0.0566)	(-0.396)	(-0.135)
Enumerator Senegalese citizen (d)	0.339	0.295	0.257	0.205
	(0.127)	(0.0996)	(0.0943)	(0.0807)
Enumerator female (d)	0.242	-0.476*	0.239	-0.0295
	(0.0848)	(-0.155)	(0.0875)	(-0.0115)
Country				
Host country = Mauritania (d)				0.709***
				(0.275)
Host country = Italy (d)				-0.223
				(-0.0853)
Constant	-1.540***	-1.492**	-0.705	-1.343***
Observations	204	163	192	559

Probit regressions: z–scores displayed, marginal effects in parentheses (evaluated at mean of independent variables). Robust standard errors.
Income, wealth and transfer variables are standardized by country to allow for different living standards.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3 Discrepancies: Evidence and interpretation

3.1 Do matched data show discrepancies between migrants' and households' reports?

Having access to a "matched" dataset, we are able to measure to what extent migrants (m) and their households of origin (h) agree on the latter's characteristics (X_h). We thus compute: $\Delta \equiv X_{hm} - X_{hh}$. Interpretation is straightforward: A positive (negative) Δ indicates that m over– (under–)estimates or states X_h.

Unfortunately, MIDDAS and migration–PSF share only a handful of variables. We can confront what m reports about h with what h themselves report for a few demographic,[20] asset ownership[21] and housing characteristics,[22] as well as the number of remitters to h.[23]

Now, the semi–qualitative survey indicated that some of those variables are more relevant in the information asymmetry context than others. 12 and 6 out of 20 respondents did not know how to explain discrepancies in demographic variables[24] and in the number of remitters, as against 1 and 2 for asset ownership and housing quality variables, respectively. Accordingly, the empirics below focus exclusively on the latter two categories, and resort to the others only when they can act as checks or confounds.

Tables 5 and 6 display the Δ's for a number of asset ownership and housing quality variables that are matched in the data. We regress Δ on a constant to obtain the mean

[20] Demographic characteristic variables are: number of people in h, the age of h's head, whether h's head is a woman, the number of h's members below 15, the number of members aged 65 or above, the number of active males and females (15–64).

[21] The asset variables refer to the numbers of: refrigerators, freezers, TV sets, DVD players, radios, CD players, fans, plows, seed drills, hoes, cars, mopeds or scooters, motorbikes, bicycles, carts and canoes.

[22] Housing variables are house type, roof material, walls material, whether h owns their dwelling, and the source of lighting and water supply. Except for the number of rooms, which is continuous (albeit left–censored), housing variables are all categorical or binary. Their modalities are ordered (from poorest to best) according to material quality, and described in Appendix Table 24 for those that are referred to in the text.

[23] A further discrepancy that can be (albeit roughly) measured is that in transfer reports. Since it is only available for the French sample, it shall be used in Section 3.2 as a robustness check rather than for its own sake. The number of remitters variable suffers from a similar issue, and shall be left aside. Finally, the type of h's phone access (none, landline, mobile phone...) shall be used in Section 5.1 as a proxy for monitoring.

[24] The Δ's for demographic characteristics are sometimes significant. The respondents who did come up with an explanation for this would blame them on migrants who "couldn't be bothered inquiring after their relatives back home"; or hypothesize an overestimation due to the fact that "when you visit the village, everyone comes to greet you".

discrepancy between X_{hm} and X_{hh}, as well as confidence intervals.[25] To allow for differential information asymmetry between remote Europe and Mauritania, which lies just across the border, we control with a Mauritanian (European) dummy when estimating the European (Mauritanian) mean. An F–test of the equality of the two means and the corresponding p–value are reported at the bottom of the table.

The asset discrepancies (Table 5) are on average very significant, and usually positive, especially in the European sample.[26] Contrary to the previous set of variables, housing characteristics display similar patterns for the European and Mauritanian samples (Table 6): In both cases housing quality[27] is significantly overestimated by the migrant.[28]

3.2 Are the discrepancies systematic?

We have just shown that discrepancies between m's and h's responses to the same survey question are usually significant, and tend to go in the same direction. This overestimation pattern is particularly noticeable for the European sample. However, non–zero Δ's (on average) need not stem from information asymmetry. Competing causes need to be confronted in order to establish that this is the most likely scenario.

Alternative causes worth investigating are: measurement error, measurement bias, the time gap between migrant (MIDDAS) and household (migration–PSF) surveys, delay in information transmission, and social desirability.[29]

Now, measurement error—provided it is "classical", as would follow from random

[25] Because of the importance of statistical significance in the empirical part of this paper, robust standard errors are used throughout.

[26] Senegalese migrants in Europe tend to overestimate all but a few (mostly, agricultural) items and TV sets and hoes, which are underestimated. Similarly for m living in Mauritania, overestimation prevails: The same assets are overestimated as in Europe, except for TV sets, radios, CD players, cars, mopeds/scooters, bicycles and canoes. Carts are overestimated, while they exhibit an insignificant discrepancy in the European sample.

[27] The number of rooms in h's house, the quality of its building and roof materials, and whether h owns it are kept as housing quality indicators as the other matched housing variables display almost no variance.

[28] A legitimate concern is that the significance of the Δ's in Tables 5 and 6 is driven by outliers. Seror (2012) shows with similar tables that the Δ's are robust to removing outliers.

[29] Of course, other factors presumably have an impact on the measured discrepancies, *e.g.* the migrant's education (as a proxy for memory or the ability to retrieve information and keep accurate accounts), and the time elapsed since she first emigrated from Senegal or since her last visit to her household of origin, etc. However, if they merely reflect communication problems, they boil down to (more serious) measurement error. Otherwise, they proxy for m's monitoring of h's behavior and actual needs, and as such they cannot act as a confound of the information asymmetry hypothesis: As will be made explicit with the model in Section 4, migrants faced with misrepresentations may engage in strategies to control what use h makes of remittances.

Table 5: Discrepancies between migrant and household reports of asset ownership ($X_{hm} - X_{hh}$)

	(1) fridge	(2) freezer	(3) tvset	(4) dvd	(5) radio	(6) cd	(7) fan	(8) plow
Mauritanian mean	0.126***	0.133***	-0.0182	0.394***	0.127	0.0727	-0.0964	0.0366
	(0.0443)	(0.0406)	(0.116)	(0.0636)	(0.130)	(0.0858)	(0.123)	(0.0440)
European mean	0.238***	0.305***	-0.430***	0.353***	0.344*	0.497***	-0.0662	0.0492
	(0.0669)	(0.0571)	(0.122)	(0.0951)	(0.198)	(0.101)	(0.253)	(0.0556)
Observations	318	317	316	315	317	314	317	286
Mauritanian mean = European mean	1.972	6.030	5.989	0.126	0.844	10.17	0.0115	0.0315
p–value	0.161	0.0146	0.0149	0.723	0.359	0.00157	0.915	0.859

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Discrepancies between migrants' and households' reports of asset ownership ($X_{hm} - X_{hh}$) (continued)

	(1) seeddrill	(2) hoe	(3) car	(4) mopedscooter	(5) motorbike	(6) bike	(7) cart	(8) canoe
Mauritanian mean	0.00599	-0.0599	0.0301	-0.0181	0.133***	0.127**	0.139***	-0.0424
	(0.0334)	(0.234)	(0.0346)	(0.0234)	(0.0408)	(0.0584)	(0.0522)	(0.0314)
European mean	0.0464	-0.620**	0.267***	0.212***	0.126**	0.722***	0.00662	0.0267**
	(0.0482)	(0.266)	(0.0660)	(0.0719)	(0.0542)	(0.138)	(0.0465)	(0.0132)
Observations	318	317	316	317	316	316	316	315
Mauritanian mean = European mean	0.473	2.496	10.08	9.249	0.0122	15.78	3.605	4.111
p–value	0.492	0.115	0.00165	0.00255	0.912	0.0000884	0.0585	0.0435

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Discrepancies between migrants' and households' reports of housing characteristics ($X_{hm} - X_{hh}$)

	(1) House type (quality)	(2) Roof material (quality)	(3) Nb of rooms	(4) h owns their dwelling
Mauritanian mean	0.200**	0.165**	1.636**	0.0952***
	(0.0856)	(0.0765)	(0.649)	(0.0270)
European mean	0.201*	0.146*	2.158***	0.0993***
	(0.115)	(0.0769)	(0.479)	(0.0262)
Observations	314	308	304	319
Mauritanian mean = European mean	0.0000880	0.0300	0.419	0.0119
p–value	0.993	0.863	0.518	0.913

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

mistakes, misunderstandings, communication problems or lapses of memory—is no cause for concern here because it should translate into wider confidence intervals and thus cannot explain how the observed discrepancies can, on average, be systematically and significantly different from zero.

We now check whether the discrepancies are robust to other interpretations than information asymmetry. We use proxies. An alternative interpretation is vindicated if the proxy is significant while the mean discrepancy loses significance.[30]

Inconsistencies in the delimitation of household boundaries The enumerators working on migration–PSF had to administer a lengthy survey to each and every constituent "cell" in h. Conversely, MIDDAS enumerators just had to ask m about a few characteristics of h. Subsequently, the (mostly) positive discrepancies found in durable assets and other variables could be the by–product of measurement bias if m tended to include relatives living next door as members of h, or if migration–PSF enumerators were prone to leave out cells if they were too many, *i.e.* if the h identified through migration–PSF is systematically a subset of that identified by MIDDAS.

Tables 7 and 8 regress asset and housing variable discrepancies that are significant for the Mauritanian and European samples, respectively, on the discrepancy in h's size,

[30]To make the tables more readily comprehensible, only the discrepancies that were found significant in Tables 5 and 6, for the Mauritanian or European samples, shall be used as dependent variables in this section.

and a constant.

Table 7: Are discrepancies driven by measurement error in household size? (Mauritanian sample)

	(1) fridge	(2) freezer	(3) dvd	(4) motorbike	(5) bike	(6) cart
Mean	0.113**	0.124***	0.402***	0.143***	0.152**	0.161***
	(0.0453)	(0.0402)	(0.0656)	(0.0441)	(0.0608)	(0.0588)
Δ h size	0.00327	-0.0105**	0.00628	0.00692	0.0212*	0.0173
	(0.00561)	(0.00437)	(0.00987)	(0.00658)	(0.0126)	(0.0115)
Observations	311	310	308	309	309	309

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Are discrepancies driven by measurement error in household size? (Mauritanian sample) (continued)

	(1) House type (quality)	(2) Roof material (quality)	(3) Nb of rooms	(4) *h* owns their dwelling
Mean	0.205**	0.158**	1.946***	0.110***
	(0.0874)	(0.0784)	(0.663)	(0.0279)
Δ h size	0.0164*	-0.00350	0.292***	0.00639*
	(0.00898)	(0.0101)	(0.0779)	(0.00334)
Observations	307	301	299	312

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Indeed, the ΔX_h's are often positively and significantly correlated with the discrepancy in the size of h. However, the mean ΔX_h's are almost never trumped by the inclusion of that variable. This implies that although some measurement bias cannot be ruled out, the Δ of h's size captures it and does not explain the ΔX_h pattern.[31]

[31] Another explanation is that ΔX_h is a function of household size. For instance, the bigger the household, the easier it is for h to misrepresent to m, blaming others or feigning ignorance when misrepresentations are exposed. Yet another is that there could be information asymmetry about household size itself, which would be positively correlated with that for other X_h. Only 2 out of 20 respondents in the semi–qualitative survey clearly explained the positive Δ in household size as h lying to m to extract more transfers. Following the example provided by one of them, it is in h's interest to exaggerate its size to m lest she thinks the rice bags she sends last longer.

Table 8: Are discrepancies driven by measurement error in household size? (European sample)

	(1) fridge	(2) freezer	(3) tvset	(4) dvd	(5) radio	(6) cd
Mean	0.238***	0.285***	-0.522***	0.339***	0.314	0.464***
	(0.0693)	(0.0601)	(0.125)	(0.0961)	(0.197)	(0.105)
Δ h size	0.00343	0.00996*	0.0400***	0.0114	0.0467**	0.0146
	(0.00657)	(0.00578)	(0.00970)	(0.00824)	(0.0186)	(0.00961)
Observations	311	310	309	308	310	307

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Are discrepancies driven by measurement error in household size? (European sample) (continued)

	(1) hoe	(2) car	(3) mopedscooter	(4) motorbike	(5) bike	(6) canoe
Mean	-0.651**	0.250***	0.207***	0.153**	0.674***	0.0238**
	(0.278)	(0.0631)	(0.0759)	(0.0607)	(0.132)	(0.0118)
Δ h size	0.0485*	0.00814	0.00514	-0.00780	0.0282**	0.00182
	(0.0251)	(0.00748)	(0.00684)	(0.00627)	(0.0131)	(0.00121)
Observations	310	309	310	309	309	308

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Are discrepancies driven by measurement error in household size? (European sample) (continued)

	(1) House type (quality)	(2) Roof material (quality)	(3) Nb of rooms	(4) *h* owns their dwelling
Mean	0.225**	0.146*	1.811***	0.0905***
	(0.114)	(0.0831)	(0.462)	(0.0256)
Δ h size	-0.00465	-0.00887	0.165***	0.00221
	(0.0119)	(0.00807)	(0.0426)	(0.00179)
Observations	307	301	299	312

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Measurement bias could also obtain if h assumed a more restrictive definition of "h's ownership of X_h" than m, driving ΔX_h upwards without information asymmetry or any exaggeration on m's part.[32] We shall use as a robustness check transfer discrepancies, which raise a similar issue. In effect, whereas in MIDDAS transfers seem to include the investment funds m sends to h, migration–PSF only includes remittances the final consumer of which is h. Unfortunately, we have no data in MIDDAS on whether m sends investment funds to h. But showing (in the paragraph on "Social desirability") that Δt is not significantly positive and not significantly correlated with the ΔX_h's shall help confirm that these are not driven by a measurement bias.

MIDDAS–PSF time gap For logistical reasons it was impossible to carry out the MIDDAS and migration–PSF surveys simultaneously. Subsequently, we do not observe X_{hm_τ} and X_{hh_τ} but X_{hm_τ} and $X_{hh_{\tau+\epsilon}}$, where τ is a time subscript and $\epsilon > 0$.[33] A legitimate question is therefore whether the observed discrepancies (Δ) are not driven by the evolution of characteristic X_h between τ and $\tau + \epsilon$.[34]

Fortunately, it can be argued that, conditional on m's place of residence and the region of Senegal where h lives, the time elapsed between MIDDAS and PSF is a random variable. In effect, households of origin of, say, Senegalese migrants in France were contacted first because MIDDAS–France was carried out before MIDDAS–Italy and MIDDAS–Mauritania. Then, the surveyors grouped the households to be contacted by regions in order to minimize transportation costs. However, households that were difficult to contact or whose addresses were ambiguous were *not* granted a particular treatment. For instance, they were not surveyed last. Subsequently, we are confident that the MIDDAS–PSF time gap is exogenous and can help us capture systematic trends in X_h among the matched sample.[35]

[32] However, the robustly negative Δ on TV set ownership and a few other durable assets provides suggestive evidence against this.

[33] The average time gap between the two surveys was four, eight and five months for France, Italy and Mauritania, respectively.

[34] Although there is no evidence of dramatic impoverishment in Senegal during those few months (see Table 11), and our variables of interest (especially housing characteristics) should not have experiencied such abrupt changes in short periods of time, the concern remains that the particular subpopulation captured in PSF did experience shocks that led them to alienate assets, yielding $X_{hm} > X_{hh}$ (since MIDDAS was implemented first).

[35] The variable we use in the regressions was built by standardizing the number of days between the migrant and household surveys over regional pairs in the host and home countries. For instance, the French sample was divided into six distinct regions, and Senegal into 12. This is meant to purge the time gap variable of

As can be seen from Tables 9 and 10, the time gap between surveys is never significant, except for refrigerators and bicycles (in the Mauritanian sample). The mean discrepancy never loses significance.

Table 9: Discrepancies between migrants' and households' reports explained by MIDDAS–PSF time gap? (Mauritanian sample)

	(1) fridge	(2) freezer	(3) dvd	(4) motorbike	(5) bike	(6) cart
Mean	0.125***	0.131***	0.366***	0.0920**	0.128**	0.118**
	(0.0470)	(0.0417)	(0.0693)	(0.0415)	(0.0634)	(0.0565)
Standardized time gap	0.121**	-0.0351	-0.0669	-0.0927	-0.188**	0.0139
	(0.0487)	(0.0411)	(0.0728)	(0.0653)	(0.0947)	(0.0824)
Observations	260	259	257	258	258	257

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Discrepancies between migrants' and households' reports explained by MIDDAS–PSF time gap? (Mauritanian sample) (continued)

	(1) House type (quality)	(2) Roof material (quality)	(3) Nb of rooms	(4) *h* owns their dwelling
Mean	0.163*	0.169**	1.669**	0.101***
	(0.0921)	(0.0802)	(0.729)	(0.0301)
Standardized time gap	-0.0725	-0.0817	-0.869	0.0190
	(0.0977)	(0.0979)	(0.789)	(0.0296)
Observations	255	253	252	260

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Delay in information transmission Although pure communication problems should not lead to discrepancies that are systematically different from zero, they could mean that information about X_h reaches m with a delay—*e.g.* income temporarily increased last year and then fell back to its original level, but since contacts are infrequent we see an overestimation today. In other words, systematic discrepancies should reflect

any correlation with geographical traits that might themselves be connected to transfers, income, etc.

Table 10: Discrepancies between migrants' and households' reports explained by MIDDAS–PSF time gap? (European sample)

	(1) fridge	(2) freezer	(3) tvset	(4) dvd	(5) radio	(6) cd
Mean	0.271***	0.328***	-0.436***	0.255**	0.169	0.494***
	(0.0755)	(0.0701)	(0.132)	(0.101)	(0.234)	(0.121)
Standardized time gap	0.140	-0.0479	0.0888	0.160	0.343	0.0540
	(0.0997)	(0.0839)	(0.157)	(0.102)	(0.330)	(0.156)
Observations	260	259	258	257	259	256

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Discrepancies between migrants' and households' reports explained by MIDDAS–PSF time gap? (European sample) (continued)

	(1) hoe	(2) car	(3) mopedscooter	(4) motorbike	(5) bike	(6) canoe
Mean	-0.799***	0.187***	0.246***	0.134**	0.704***	0.00896
	(0.264)	(0.0712)	(0.0863)	(0.0664)	(0.159)	(0.00897)
Standardized time gap	-0.815	0.0162	0.0960	-0.0516	-0.205	-0.00568
	(0.584)	(0.0746)	(0.0722)	(0.0628)	(0.211)	(0.00575)
Observations	258	258	259	258	258	257

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Discrepancies between migrants' and households' reports explained by MIDDAS–PSF time gap? (European sample) (continued)

	(1) House type (quality)	(2) Roof material (quality)	(3) Nb of rooms	(4) *h* owns their dwelling
Mean	0.203	0.162*	1.853***	0.0882***
	(0.140)	(0.0886)	(0.548)	(0.0297)
Standardized time gap	0.118	0.139	-0.0325	0.0148
	(0.147)	(0.104)	(0.593)	(0.0297)
Observations	255	253	252	260

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

economic shocks in *previous* periods. Once again, it is unlikely that for a majority of migrants who remit regularly (see Tables 1 through 3)[36] contacts should be so infrequent that sharp changes in living standards could go unnoticed, and Senegal experienced no recession in the past years, GDP per capita growing steadily between 2005 and 2011 (Table 11). But it is quite conceivable that emigration households' economic conditions do not follow the national patterns.

Table 11: Senegalese Gross domestic product per capita in purchasing power parity dollars (current international dollars), 2005–2011

Year	GDP
2005	1,565.479
2006	1,617.107
2007	1,706.224
2008	1,757.797
2009	1,770.311
2010	1,819.222
2011	1,877.851

Source: World Economic Outlook Database (April 2011).

However, economic shock variables are generally unable to explain the observed discrepancies, *i.e.* either the economic shock variable is insignificant, or it is significant but the mean discrepancy (controlling for the shock) remains significant and does not change sign. Due to limited space and because we cannot guarantee that the economic shocks used to test the "Delay in information transmission" hypothesis are exogenous to h's past actions and m's remittances, estimates are not reported. The interested reader should refer to Seror (2012).

Social desirability effect Another serious confound is that the Δ's could be inflated by a "social desirability effect": m (h) might overdo h's (their own) wealth to "show off" her extraction or the beneficial effect of remittances. Conversely, h (or m) might pass themselves (h) off as poorer than they actually are so as to benefit from financial assistance—or because they think that is what the enumerator wants to hear. The direction of the social desirability bias is thus *a priori* indeterminate. However, it is unlikely that h could lie to the enumerator on such variables as assets or housing quality, all

[36] Migrants also report on average weekly or semimonthly phone calls with the household of origin.

easily verifiable while implementing the survey in h's home.[37]

In order to check for the role played by desirability bias, let us first use enumerator characteristics as proxies, and then consider a general tendency to exaggerate thanks to transfer report discrepancies.

A nice feature of the MIDDAS survey can help us study empirically whether the migrant was showing herself in a favorable light. We use a binary variable that equals 1 for female enumerators to capture this effect, as (predominantly male) migrants may try to "show off" in front of a woman. This dummy is arguably exogenous as the enumerators hired for MIDDAS were randomly allocated to survey places.

Tables 12 and 13 show whether female enumerators explain the discrepancies significant in the Mauritanian and European samples, respectively. The social desirability effect is sometimes strong enough to exhaust the observed gaps between migrants' and households' reports, suggesting that the enumerator variable "trumps" other factors underlying the discrepancy.

Table 12: Discrepancies between migrants' and households' reports explained by the enumerator's gender? (Mauritanian sample)

	(1) fridge	(2) freezer	(3) dvd	(4) motorbike	(5) bike	(6) cart
Mean	0.142***	0.0787*	0.317***	0.00794	0.0397	0.0476
	(0.0430)	(0.0413)	(0.0674)	(0.0309)	(0.0429)	(0.0464)
Enumerator female	-0.0667	0.229**	0.324*	0.531***	0.371*	0.388**
	(0.132)	(0.112)	(0.167)	(0.124)	(0.204)	(0.161)
Observations	318	317	315	316	316	316

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

By the same token, Senegalese surveyors could elicit exaggerated Δ's as there may be a "competition between migrants, each one trying to remit more than the other", in the words of a migrant interviewed in October 2012. This too can be shown in a

[37] Only one respondent in the semi–qualitative survey speculated that "the household believed the enumerator could give them money". Besides, we expect m's statements to be less subject to a desirability bias if she accepted that her household of origin be surveyed, as she would have known that the information she provided could be checked in Senegal.

Table 12: Discrepancies between migrants' and households' reports explained by the enumerator's gender? (Mauritanian sample) (continued)

	(1) House type (quality)	(2) Roof material (quality)	(3) Nb of rooms	(4) *h* owns their dwelling
Mean	0.136	0.175**	1.468*	0.0938***
	(0.0962)	(0.0783)	(0.791)	(0.0282)
Enumerator female	0.264	-0.0430	0.711	0.00625
	(0.208)	(0.220)	(1.297)	(0.0749)
Observations	314	308	304	319

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Discrepancies between migrants' and households' reports explained by the enumerator's gender? (European sample)

	(1) fridge	(2) freezer	(3) tvset	(4) dvd	(5) radio	(6) cd
Mean	0.164	0.200*	-0.618**	0.759***	0.873**	0.660***
	(0.103)	(0.105)	(0.253)	(0.192)	(0.410)	(0.211)
Enumerator female	0.118	0.165	0.295	-0.634***	-0.831*	-0.254
	(0.135)	(0.124)	(0.283)	(0.214)	(0.456)	(0.236)
Observations	318	317	316	315	317	314

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Discrepancies between migrants' and households' reports explained by the enumerator's gender? (European sample) (continued)

	(1) hoe	(2) car	(3) mopedscooter	(4) motorbike	(5) bike	(6) canoe
Mean	-0.463	0.182	0.273**	0.0909	0.673***	1.36e-17
	(0.300)	(0.127)	(0.125)	(0.0869)	(0.153)	(.)
Enumerator female	-0.245	0.134	-0.0956	0.0549	0.0773	0.0421**
	(0.486)	(0.147)	(0.153)	(0.111)	(0.251)	(0.0207)
Observations	317	316	317	316	316	315

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Discrepancies between migrants' and households' reports explained by the enumerator's gender? (European sample) (continued)

	(1) House type (quality)	(2) Roof material (quality)	(3) Nb of rooms	(4) h owns their dwelling
Mean	0.0185	0.173*	0.654	0.127***
	(0.174)	(0.101)	(0.547)	(0.0452)
Enumerator female	0.287	-0.0426	2.404***	-0.0439
	(0.230)	(0.147)	(0.869)	(0.0554)
Observations	314	308	304	319

Standard errors in parentheses

The whole sample is used; country dummies and interactions are not displayed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

quasi–experimental manner thanks to an enumerator citizenship dummy. The interested reader is invited to consult Seror (2012) for the results based on the enumerator's citizenship.[38]

To conclude, the surveyor's citizenship and gender do play a role, but most discrepancies remain robust in sign and significance when enumerators' characteristics are controlled for.

The semi–qualitative survey showed that in total (*i.e.* at least for one of the following categories of discrepancies: assets, housing quality or number of remitters) 7 out of 20 respondents deemed plausible that m exaggerated her role in improving h's living standard, and subsequently X_h. This tendency is arguably captured by the discrepancy in tranfers according to m's and h's reports. For the French sample, we can spot in migration–PSF whether a particular transfer comes from the migrant surveyed in France or not.[39] This means we can directly compare transfers reported by m and by h.

[38]The gender dummy provides a better test of the social desirability hypothesis. In effect, although both variables are arguably exogenous, they may influence the Δ's through two channels. On the one hand, we hypothesized that in both cases Δ would tend to be larger. But on the other hand, Senegalese surveyors were more successful in convincing fellow migrants to put them in touch with h, whereas it was the opposite for female enumerators. Now, if the migrants in the matched sample were those who had *least* exaggerated X_h, we have a negative (positive) bias in the regression of ΔX_h on the dummy for whether the enumerator was Senegalese (a woman). Subsequently, we are more confident in rejecting the hypothesis that the coefficient is positive in the regressions focusing on gender. Note however that regressing a dummy equal to 1 when h could be surveyed on the X_{hm}'s does not yield a consistent pattern (not reported).

[39]It is impossible to tell which transfer was received from m in the reports of the households of origin of migrants living in Italy and Mauritania.

However, it is impossible to distinguish between missing and zero values in migration–PSF: m will not appear in the transfers received by h either because m sent nothing to h in the past twelve months or because h forgot to mention her (or the answer was poorly recorded). Therefore, two variables are constructed for the discrepancy in total transfers from migrants living in France ($\Delta t \equiv t_{mm} - t_{mh}$): The first one considers that if m is absent from h's transfer reports it is a missing value; the second treats m's absence as evidence of an absence of transfers ($t_{mh} = 0$), which increases sample size and mechanically inflates Δt.

Because sample size is severely reduced (between 45 and 75 observations, depending on the definition of t), we study Δt graphically.[40] Although no confidence intervals can be provided, kernel density estimations (Figure 1) of the two variables and for non–zero values suggest that some overreporting from m cannot be ruled out, especially at the bottom of the distribution, which is consistent with some exaggeration.[41]

As can be seen from Table 14, the discrepancy in m's transfers to h is almost never a significant predictor of discrepancies in assets and housing quality. Although sample size is quite limited, this seems mostly driven by very small point estimates with reasonable standard errors.

Admittedly, it could still be argued that m overreports X_h intentionally. But the fact that Δt is seldom significant and almost never trumps the mean ΔX_h suggests that social desirability (or m's tendency to "exaggerate") and a measurement bias in the way m and h delimit what belongs to h are not driving the results.

3.3 Is the rent–extraction hypothesis positively supported by the data?

We have so far used a process of elimination. Even if the data seem to support the information asymmetry hypothesis, "positive" empirical evidence that rent–extraction is likely to drive most non–zero Δ's can be provided.

[40] Regressions reproducing for Δt what Tables 5 and 6 did for X_h are available upon request: Δt is only significant when missing values are coded as 0.

[41] The next step is to see what explains the (sometimes) significant discrepancies in the unreported Δt regressions in order to disentangle exaggeration from other explanations. The time gap between the two surveys does not explain the discrepancies in transfers (it is either insignificant or negative); so a "crisis" in migrants' remittance capacity between MIDDAS and migration–PSF cannot blamed. Regressions of Δt on the enumerator characteristics introduced above (also available upon request) show that when the enumerator dummies are significant they are large and exhibit the expected sign, which seems to suggest that social desirability is at work.

Table 14: Are discrepancies correlated with the overreporting of transfers? (French sample)

	(1) freezer	(2) freezer	(3) tvset	(4) tvset	(5) cd	(6) cd	(7) car	(8) car	(9) bike	(10) bike
Δt (€1000)	0.0373*		0.0135		0.00826		-0.00841		0.0895	
	(0.0210)		(0.0490)		(0.0303)		(0.0282)		(0.0831)	
Mean	0.211**	0.186**	-0.683***	-0.580***	0.278*	0.256**	0.162*	0.0993	0.739***	0.546***
	(0.0860)	(0.0796)	(0.159)	(0.137)	(0.139)	(0.119)	(0.0841)	(0.0684)	(0.227)	(0.193)
Δt (missing to 0, €1000)		0.0169		0.00871		-0.00356		-0.00742		0.0677
		(0.0241)		(0.0401)		(0.0272)		(0.0239)		(0.0663)
Observations	50	75	50	75	50	75	50	75	50	75

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Are discrepancies correlated with the overreporting of transfers? (French sample) (continued)

	(1) Roof material (quality)	(2) Roof material (quality)	(3) Nb of rooms	(4) Nb of rooms	(5) *h* owns their dwelling	(6) *h* owns their dwelling
Δt (€1000)	0.0411		0.266		-0.0128	
	(0.0535)		(0.435)		(0.0188)	
Mean	0.234	0.262*	3.276***	2.982***	0.103**	0.102***
	(0.168)	(0.143)	(0.932)	(0.712)	(0.0440)	(0.0384)
Δt (missing to 0, €1000)		0.0247		0.208		0.00574
		(0.0450)		(0.319)		(0.0186)
Observations	48	71	44	68	50	75

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

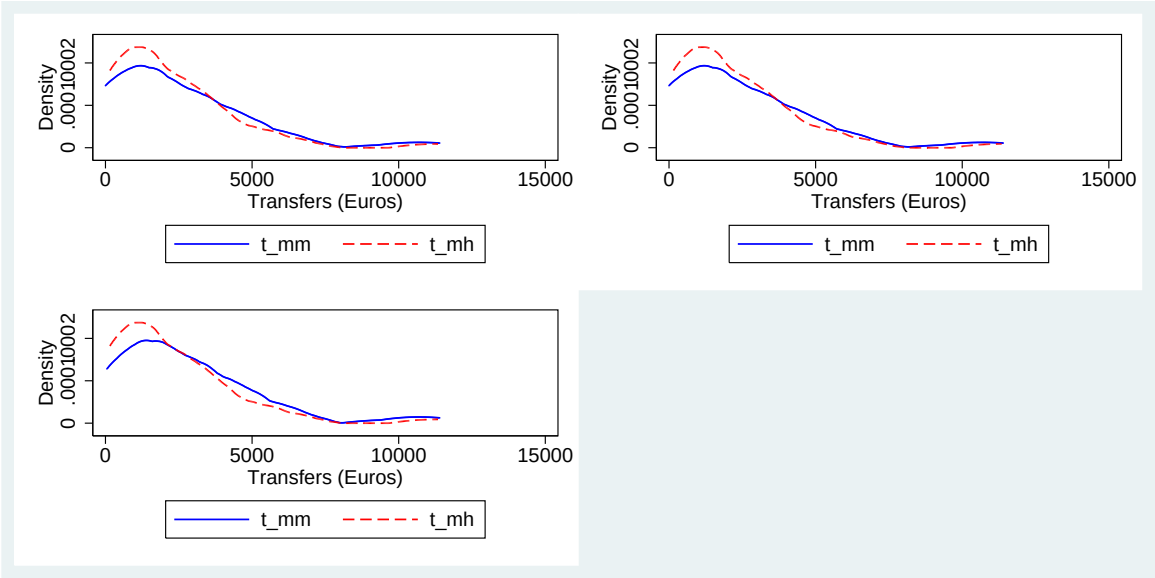

Figure 1: Kernel density plots of t_{mm} and t_{mh} (French sample)

The geographical gradient of discrepancies Since Tables 5 and 6, we have seen that the Δ's are consistently smaller and less often significant in the Mauritanian than in the European sample. This is quite intuitive: Migrants in Mauritania are just a few hours of car or bus away from their households of origin and can easily travel back and forth, which is expected to reduce gaps between m's and h's reports. Such geographical proximity and frequency of contacts should reduce the noise due to communication difficulties and delayed information transmission. But it may also explain the fact that ΔX_h is systematically non–zero if closer links between m and h reduce the room for information asymmetry.

Indeed, frequency of contact and other monitoring proxies are significantly correlated with the Mauritanian dummy. The dependent variable considered in Table 15 is the frequency of m's visits to h (in times per year, 0 meaning that m never went back) but the time elapsed since m's last visit (in days) and a dummy for whether m remits in kind, which can be argued a way to impose a consumption pattern (and thus to monitor) and is more convenient if distances are small or transportation inexpensive, give similar results (not reported). The regression controls for whether m has a spouse or children living with h; this is necessary insofar as closer links could be due to altruism

Table 15: Discrepancies smaller when m is in Mauritania than in Europe

	(1) Nb of visits to *h* per year	(2) Nb of visits to *h* per year	(3) Nb of visits to *h* per year	(4) Nb of visits to *h* per year	(5) Nb of visits to *h* per year
m lives in Mauritania	0.650***	0.650***	0.612***	0.636***	0.645***
	(0.118)	(0.117)	(0.119)	(0.118)	(0.118)
Spouse lives with *h*		0.442***			0.336**
		(0.135)			(0.168)
At least one child lives with *h*			0.345***		
			(0.121)		
Nb of *m*'s children living with *h*				0.0901***	0.0444
				(0.0310)	(0.0385)
Constant	0.820***	0.724***	0.719***	0.740***	0.705***
	(0.0693)	(0.0754)	(0.0785)	(0.0753)	(0.0774)
Observations	895	892	880	880	879

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and not to monitoring.[42] Since the estimated coefficient is very robust to the introduction of such proxies for altruism, Table 15 lends support to the information–asymmetry interpretation of the discrepancies.

Age of h's head: A by–product of rent–extraction It is unlikely that all significant discrepancies deserve the same interpretation. For instance, it is not obvious why (and even *how*) h would (could) delude m into thinking that its head is older than he or she actually is, unless m thinks the head is another, older person. Now, it can be seen from Table 16, Column 1, that m overestimates the age of h's head by 3.46 years on average, which is significant at any conventional confidence level.

A natural rationale would be that m can be systematically wrong (on average) about the head's age if a younger relative is actually in charge of the household and lets m believe that the patriarch is still at the helm so as to ensure m's loyalty and thus secure remittance flows.[43]

However, the discrepancy in the age of h's head might capture a measurement bias due to differences in enumerators' training for MIDDAS and migration–PSF. The surveyors hired for the latter had worked on the original PSF project, for which a lot of effort was had been put into defining household boundaries and decisionmaking clearly to the enumerators. Subsequently, MIDDAS data might systematically overestimate the age of h's head as the migrant would spontaneously name the "patriarch", whereas in migration–PSF better–trained enumerators made sure h named who is *actually* at the helm. Now, the semi–qualitative survey comprised a question on this, and it turns out that 18 out 20 respondents think that the head they mentioned *both* runs h nominally *and* makes everyday decisions, which goes counter to the measurement–bias theory.[44]

A stronger rejection comes from Table 16: The discrepancy is much higher in Europe than in Mauritania, where it is merely significant at the 90% confidence level. This

[42] Actually, these controls are probably too "strict" because m could also be more careful in her monitoring of h if her spouse and children live with them.

[43] Seror (2012) shows that other potential rationales for m's mistakenly reporting an older h's head, namely the MIDDAS–PSF time gap and news (of his or her death) taking time to reach m, had no explanatory power when the discrepancy in the age of h's head was the dependent variable. The discrepancy cannot be blamed either on household boundary delimitation problems or social desirability (regressions not reported).

[44] A caveat is that the sample might be quite different from the MIDDAS one. For instance, if MIDDAS respondents tended to belong to an "older" generation, they might have answered in deference to the "patriarch" while knowing perfectly well that he or she does not actually run h.

suggests that even if some measurement bias cannot be ruled out, it cannot explain this geographical differential. Moreover, when one controls for h's room for rent extraction, we see that the time since m's last visit enters the regression very significantly for the Mauritanian sample, has the expected sign and makes the discrepancy vanish into thin air.

Table 16: m tends to overestimate the age of h's head

	(1) Δ age of h's head	(2) Δ age of h's head	(3) Δ age of h's head
Mean	3.461***		
	(0.923)		
Mauritania mean		2.012*	-1.167
		(1.113)	(1.416)
European mean		5.113***	4.126*
		(1.506)	(2.168)
Years since last visit to h (Mauritania)			1.908***
			(0.619)
Years since last visit to h (Europe)			0.428
			(0.720)
Observations	304	304	256

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17 shows that the discrepancy in the age of h's head is entirely explained by m's being mistaken about his or her identity. Although this piece of evidence is available only for the matched French sample, it suggests that this discrepancy is a means to make rent extraction possible rather than information manipulation aimed to have a direct impact on received transfers.

Interestingly, the discrepancy in the age of h's head is a significant predictor of, and has the same sign as several asset and most housing variable discrepancies. This is evidence that misrepresentation about the age of h's head proxies h's very ability to misrepresent X_h (regressions not reported).[45]

[45]Misrepresentations about the gender of h's head, while less significant in general, give similar results as those about his or her age, yielding consistently coefficients of the opposite sign (since m tends to *un-der*estimate how often h's head is a woman) when used to predict other discrepancies.

Table 17: Is the migrant mistaken about who actually runs the household of origin?

	(1) Δ age of h's head
Head of household of origin same in m's and h's reports	-11.34***
	(4.202)
Mean	10.76***
	(4.031)
Observations	80
Mean + Same head = 0	0.241
p value	0.625

Standard errors in parentheses

The regressions are run on the matched sample for Senegalese migrants in France.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Determinants of information asymmetries Another way to establish that information asymmetry is the main explanation for systematically non–zero Δ's is to look at the explanatory power of variables that reflect situations conducive to rent extraction.[46] Since we cannot argue the estimates to be causal, we shall not report them here. They however give credence to the rent–extraction hypothesis.[47]

Further support for the rent–extraction hypothesis is yet needed, which will be provided in Section 5.

[46]The Δ's can be aggregated into indices that simply sum across standardized Δ's within sets of X_h—asset and housing variables—and used as dependent variables. The regressors included are the following. We introduce the educational gap between m and h's head to capture the idea that an educational advantage can help detect misrepresentations and thwart rent extraction. The number of emigrants from h (who do not reside with m) represent h's outside options: If misrepresentations are detected by m, h can maintain their living standards through temporary retaliations thanks to other remittance sources. The expected sign is thus positive. Finally, the existence where m lives of an association of migrants from the same village or neighborhood in Senegal is taken as a proxy for the density of the migrant population and strength of the remittance norm. This interpretation follows from the semi–qualitative survey, where 8 out of 20 respondents said that "living in a community with many migrants from the same neighborhood or village in Senegal increases the pressure to send money back home". If significant, this variable should enter positively. m's total income, h's wealth score (m's report), a European dummy, and indicator variables for whether m is h's head's son or daughter, and whether she has a spouse or children living with h are also included to deal with endogeneity.

[47]The only significant regressors (when assets are considered, not housing variables) are the European dummy and the indicator variable for the presence of a migrants' association where m lives. They both have the expected sign. The "number of emigrants from h" is never significant but does approach the 90% confidence cutoff in several specifications, which implies that h may have outside options in case their distorting of information were detected by m and their remittance contract with her subsequently jeopardized.

4 Model and predictions

To identify testable predictions about m and h's behaviors under information asymmetry, one first needs a theoretical framework. This is what we present here. The model also allows us to make a few explicit assumptions necessary to derive the potential bias in transfer regressions when information asymmetry is not taken into account—see Section 6.

4.1 ΔX_h as evidence of conflicting consumption patterns

We draw on Roger's (2010, 2012) soft–information framework, adapt it to the migration context and extend its predictions to the relationship between information asymmetries, monitoring and remittances. The main novelty of our model is to swerve from the tradition of considering transfers as a function of an aggregate variable such as income or living standards. Most of the literature aims indeed at disentangling the effects of different remittance motives—altruism, purchase of services, repayment of loans to migrate, insurance, inheritance, and Stark's (1999) "strategic motive", according to Rapoport and Docquier's (2006) typology[48]—or at determining the "balance between altruism and self–interest" (Carling, 2008). But remittance motives often coexist in aggregate variables and are therefore difficult to tell apart: "Researchers have found it a greater challenge [...] to quantify the relative importance of the different motivations or to establish strong causal linkages that could rule out alternative explanations for remittance flows." (Yang, 2011)[49]

We view each of the X_h's not as proxies for income, but as consumption items that m finances through remittances and that m and h may prioritize differently. For instance, migrants earmark some of their remittances (or directly use in–kind transfers) for the purchase of a refrigerator, and h may or may not use the money according to m's plans (in the in–kind context, h can resell the good). ΔX_h is then evidence that

[48] The repayment–of–loans motive—supported, for instance, by Ilahi and Jafarey's (1999) results—does not lend itself to information distortion, except for "escape clauses allowing the migrant to default on his commitment, like illness or a spell of unemployment" (Azam and Gubert, 2004), which we shall discuss in Section 7.2.

[49] In the words of Rapoport and Docquier (2006): "[A]t a micro level, it is extremely difficult to discriminate between competing theories of remittances, which often share similar predictions as to the impact of the main explicative variables, implying that truly discriminative tests have to rely on additional variables whose details are not always available."

m's and h's consumption preferences for h need not be aligned.

This modeling is supported empirically. First, MIDDAS and migration–PSF questionnaires asked both m and h whether "according to them" h is poorer, richer than or has the same living standard as the community they live in. Now, m systematically and significantly estimates that h is *poorer* than they think themselves. The measure is subjective and rough but it seems to reject the idea that X_h is a proxy for aggregate income or living standards. Second, the model follows the answers to the semi–qualitative survey: 5 out of 20 respondents either earmark funds for particular purchases or send in–kind transfers; 7 invest in Senegal (be it in real estate or in a business), 3 of which invest through h; and 16 say their transfers are explicitly meant to benefit the whole household, not particular members. More strikingly, 19 out of 20 respondents earmark transfers in one of these ways or another.

In this context, the model views information asymmetry as rent extraction. By "rent extraction", we mean that misrepresentations are due to h manipulating information in order to maximize the transfers received from m, while trying to consume as they like. In other words, we propose a more restricted conception of remittance motives: m "invests" in or through h[50]; because this investment pertains to a conscious plan, m earmarks all or part of her transfers according to *her* consumption preferences for h, with which h may beg to differ.

m can observe the X_h's upon visiting h or through detailed phone calls (to h or other acquaintances that might be in contact with h). We refer to these actions, taken to reduce the informational gap, as "monitoring".

4.2 Soft–information framework

Except for infrequent home trips, migrants' transfer decisions are likely to be based on recipients' *reports* of their compliance with the targets rather than on verifiable, "hard" evidence. Thus, the standard moral hazard literature, where effort may be unobservable but outcomes at least are common knowledge, does not provide a suitable framework to study information asymmetries between remittance senders and recipients. Migrants'

[50] m may invest in h's well–being, which is synonymous with altruism, in real estate or any project for the realization of which she needs h's intervention, or in prestige (which can be construed as capitalizing for one's return or a way to repay one's family for past debts).

decisions are based on "soft information": Unverifiable actions *and* outcomes.[51]

Setting and timing The migrant is modeled[52] with very few instruments (a transfer scheme and an audit technology), which prevents her from disentangling *ex ante* moral hazard (effort provision to comply with m's earmarking)[53] and *ex post* adverse selection (misrepresentation) problems that arise in the soft–information context.

There are two actors: h (the agent) and m (the principal). The former (latter) is risk–averse (risk–neutral). h undertakes an action $a \in A$. a should be construed in a broad sense: Abiding by m's earmarking of the transferred monies is a "higher action". h incurs an increasing and convex cost $c(a)$. a transforms m's earmarking into X_{hh}, which falls within a range $[\underline{X_{hh}}, \overline{X_{hh}}]$. X_{hh} is the actual value of X_h. One can view X_{hh} as a stochastic outcome as h may exert some effort in the direction of m's goals and eventually lose heart, allocating the money to some more pressing need.[54] X_{hh} thus has conditional distribution $F(X_{hh}|a)$, the density of which satisfies the Monotone Likelihood Ratio Property, *i.e.* a higher a yields a stochastically dominant function $F(\cdot)$, in the first–order sense.

Only h observes X_{hh} and reports a message $X_{hm} \in [\underline{X_{hm}}, \overline{X_{hm}}] \supset [\underline{X_{hh}}, \overline{X_{hh}}]$. h's net utility is $u(t, a) = v(t) - c(a)$, where t denotes the transfers received. On the other hand, m's net payoff is $S(t, X_{hh}) = X_{hh} - t$.[55]

m sends transfers $t(X_{hm})$ [56] but can resort to an "audit technology" (akin to a sampling process), thus partially restoring *ex post* observability. The monitoring technol-

[51] For instance, Azam and Gubert (2004) explicitly assume that m can observe h's agricultural output, but not their effort, which means that geographical differentiation does not aggravate information asymmetries compared to the standard moral hazard in agricultural production framework. Similarly, in De Laat (2008) the migrant husband can observe farm output or news from the village about his wife's fidelity. Finally, in Ashraf et al. (2011) there is suggestive evidence of moral hazard, be it due to h's conscious diversion of remittances meant to be saved or to self–control problems, but none of information manipulation, *i.e.* both effort and outcomes could be observable *ex post*.

[52] We do not derive the full model here but delineate the setting, present the problem faced by the migrant, and discuss briefly the predictions.

[53] Effort can be construed as the cost to h of following m's rather than their own consumption preferences. We do not model their preferences explicitly, only the utility they derive from transfers.

[54] X_{hh} is also stochastic if effort is interpreted, as one interviewee hypothesized, as h being less careful with what they receive from m, not realizing how much it costs or how difficult it was to earn the money to purchase X_h.

[55] Note that X_{hh} being unobservable does not prevent m from maximizing $\mathbf{E}[S(t, g(X_{hh}))]$.

[56] The function $t(\cdot)$ can be increasing or decreasing in its argument, *i.e.* depending whether m wants h to have more or less of X_h. For the sake of simplicity, we consider only the case where $t'(\cdot) > 0$ in what follows. This is all the more plausible as Senegal is a poor country and therefore remittances are rather meant to increase asset ownership and housing quality. "Positive" targets are thus more realistic, and presumably: more acceptable to h, than a *reduction* in assets, etc.

ogy, α, is implemented through phone calls, physical visits, conversations with other remittance senders, etc. However, contrary to Roger (2012), the marginal cost of α—$k(\alpha)$, $k'(\alpha) > 0$, $k''(\alpha) > 0$—is positive, so that audits may not be run and yet transfers are sent. Moreover, monitoring is imperfect: It detects discrepancies with probability $0 \leq p(X_{hm} - X_{hh}; \alpha) \leq 1$, which is a function of how much m invests in monitoring and how blatant h's misrepresentation is. It is continuous and differentiable in both arguments, and $p(0; \alpha) = p(X_{hm} - X_{hh}; 0) = 0$.

If a discrepancy is discovered, h gets nothing. This simplifying assumption is grounded in survey evidence: 10 out of the 20 respondents[57] of the semi–qualitative survey were asked questions on how they would react if they realized that h distorted information to conceal a breach of the earmarking; and 7 out of those 10 said that they would cut transfers at least temporarily in retaliation.[58] Therefore, their expected utility is: $U = v(t(X_{hm}))[1 - p(X_{hm} - X_{hh}; \alpha)]$.[59]

The timing of the game is as follows:

1. m offers a contract $C = \langle [\underline{X_{hm}}, \overline{X_{hm}}], t(X_{hm}), p(X_{hm} - X_{hh}; \alpha) \rangle$. A critical assumption here is that m can commit to the contract, even when presented with an obvious exaggeration;[60]

2. if h do not reject the offer, they choose a;

3. a generates X_{hh};

4. h report X_{hm} such that $\max_{X_{hm}} v(t(X_{hm}))[1-p(X_{hm}-X_{hh})]$, *i.e.* h's optimal $X_{hm} = X_{hm}(X_{hh})$, holding α constant and using the notation $\Delta \equiv X_{hm} - X_{hh}$, is defined by:

$$v't'(X_{hm})[1 - p(\Delta)] - v(t(X_{hm}))p'(\Delta) = 0 \tag{1}$$

[57] The smaller subsample is due to screening questions. For instance, if the respondent does not invest in Senegal, the survey item regarding her reaction in case h embezzled the investment funds is irrelevant.

[58] Interestingly, one interviewee put forward that she could not cut transfers as her relatives do not have other income sources, but she would throw out the threat of putting an end to remittances.

[59] The cost of action, $c(a)$, does not show in h's expected utility because it is sunk at the time of information revelation: Whether h expended effort or not should not enter their decision to misreport.

[60] This is not particularly unrealistic in a context where remitting is a strong social norm. In the semi–qualitative survey, 12 out of 20 feel that remitting to h is a "moral obligation". Indeed, *all* interviewees declared remitting to relatives in Senegal; and 15 of them do so on a regular basis.

5. α occurs;

6. m remits according to the contract.

Three cases arise. First, there is truthful revelation if:

$$v(t(X_{hh})) \geq v(t(X_{hm}))[1-p] \tag{2}$$

So h tells the truth if X_{hh} maximizes $v(t(X_{hm}))[1-p(X_{hm}-X_{hh})]$. Using Equation 1, this implies that: $v't'(X_{hh}) \leq v(t(X_{hh}))p'(0;\alpha)$.[61] Now, if $\widetilde{X_{hh}}$ satisfies the F.O.C. (Equation 1) and lies *between* $\underline{X_{hh}}$ and $\overline{X_{hh}}$, *partial* revelation obtains. Third, there is no truthful revelation if $v't'(X_{hh}) \geq v(t(X_{hh}))p'(0;\alpha)$, $\forall X_{hh} \in [\underline{X_{hh}}, \overline{X_{hh}}]$.

The migrant's problem can be stated as follows:

$$\max_{\alpha,t,a} \int_{\underline{X_{hh}}}^{\widetilde{X_{hh}}} [\theta - (1-p(X_{hm}(\theta)-\theta;\alpha))t(X_{hm}(\theta))]\,\mathrm{d}F(\theta|a) + \int_{\widetilde{X_{hh}}}^{\overline{X_{hh}}} [\theta - t(\theta)]\,\mathrm{d}F(\theta|a) - k(\alpha) \tag{3a}$$

$$\text{st} \quad X_{hm}(X_{hh}) = \arg\max_{X_{hm}} v(t(X_{hm}))[1-p(X_{hm}-X_{hh})] \tag{3b}$$

$$\text{and} \quad \int_{\underline{X_{hh}}}^{\widetilde{X_{hh}}} v(t(X_{hm}(\theta)))[1-p(X_{hm}(\theta)-\theta)]\,\mathrm{d}F(\theta|a) + \int_{\widetilde{X_{hh}}}^{\overline{X_{hh}}} v(t(\theta))\,\mathrm{d}F(\theta|a) - c(a) \geq 0 \tag{3c}$$

$$\text{and} \quad \int_{\underline{X_{hh}}}^{\widetilde{X_{hh}}} v(t(X_{hm}(\theta)))[1-p(X_{hm}(\theta)-\theta)]\,\mathrm{d}F_a(\theta|a) + \int_{\widetilde{X_{hh}}}^{\overline{X_{hh}}} v(t(\theta))\,\mathrm{d}F_a(\theta|a) = c'(a) \tag{3d}$$

Equations 3b, 3c and 3d represent h's information revelation, the participation and the moral hazard constraints, respectively.

4.3 Predictions of the model

First, as can be seen from the model, α, the optimal a and the transfer scheme t are jointly determined, and h may optimally choose not to report the true X_{hh}.

Second, as there are no fines in the model, the participation constraint is slack: Equation 3c does not hold with equality and thus h receives an *ex ante* rent.[62] However, we can still observe $\Delta = 0$ since after a successful audit m knows the true X_{hh}. But on

[61] Intuitively, truthful revelation obtains if for all values of the private information X_{hh}, the marginal utility of transfers (and thus of misrepresenting) is below the marginal cost h would incur if they departed from the true X_{hh} to inflate remittances.

[62] Complete truth telling is possible theoretically if there are negative transfers or a penalty l for misreporting, so that Equation 2 becomes $v(t(X_{hh})) \geq [v(t(X_{hm}(X_{hh}))) - v(-l)][1-p]$. For l large enough, truthful revelation is possible.

average we should see systematic differences. These first two predictions are supported by the results in Section 3.

Third, optimal α and t are expected to co–vary positively because "the more powerful the *ex ante* incentives for high effort (*i.e.* the steeper the transfer function), the more attractive is the option to manipulate information *ex post*, especially when it is bad" (Roger, 2012). In effect, with soft information transfers need to increase with reported effort (as in standard moral hazard models, where contracts are incentivized) but this calls for investing in "audit technologies" that keep a check on the agent's information distortion. Interestingly, this prediction together with the first one implies that transfers should not be lower but higher with information asymmetry compared to the perfect–information world because of the double problem of *ex ante* moral hazard and *ex post* adverse selection: The latter requires an *ex ante* rent to provoke truth telling, while the former requires a steeper transfer function to make effort (compliance with m's instructions) more attractive.

Fourth, in case of partial revelation, discrepancies should be larger at the bottom (top) of the distribution of X_{hh} when $t'(\cdot) > 0$ (< 0) and $\Delta > 0$ (< 0)—*i.e.* when there is rent extraction. This intuitively follows from the fact that the agents with the worst private information (who move farthest away from m's instructions) have the lowest cost of misreporting and therefore the "strongest incentives to misreport when facing a concave transfer" (*ibid.*).

Last but not least, the signs of Δ and of $t'(\cdot)$ can help us determine empirically whether rent extraction indeed underlies the observed discrepancies. If the recipient expects m to increase (cut back) transfers when X_h rises, then the difference between reported and actual X_h will be positive (negative) if rent extraction is the primary motive for information distortion.[63] In another context, where instead of particular X_h's we had more aggregate variables, studying information asymmetries might help shed a new light on the debate over motives for private transfers (Rapoport and Docquier, 2006, *inter alia*). Indeed, if misrepresentation is a reasonable assumption, then the

[63] Conversely, if h is "altruistic", *i.e.* they misreport in order to receive less and alleviate the pressure on m to remit, the data should exhibit Δ's and $t'(X_{hm})$ of opposite signs; or simply if the Δ's are not correlated with the transfers, we will observe an absence of stable pattern.

signs of Δ and of $t'(\cdot)$ can help show in which direction rent extraction operates and therefore which motive dominates.[64]

The most important predictions (covariance of t and α, misrepresentation at the bottom and the sign of $t'(X_{hm})$) are investigated in Section 5. The model delineated here will also enable us to derive the bias potentially inflicted by the oversight of information asymmetry in transfer regressions (Section 6).

5 Empirical test of the model's predictions

The model provides testable predictions, *i.e.* empirical patterns that are difficult to explain outside the model, and would therefore, if supported by the data, show that rent extraction is the most likely interpretation of the observed discrepancies.

5.1 Covariance of transfers and monitoring

The data offer a number of proxies for m's monitoring activities: frequency of phone contacts[65] with and physical visits to h, frequency of contacts with other migrants who remit to h,[66] and whether m sends at least part of her remittances in kind.[67]

Unfortunately, we lack proper instruments for monitoring activities and thus regressing total transfers on them is likely to yield biased results. First, because transfers and monitoring are jointly determined by the prevalence of information problems. Second, because our proxies for "audits" also reflect other aspects of m and h's relationship or characteristics. For instance, a higher frequency of phone contacts can indeed suggest more careful monitoring, but it is also a manifestation of closer ties or that they can afford steep phone bills, which in turn might affect transfers. This is why Tables 18 and 19 control for m's total income and whether m's spouse or children live with h.

We also use proxies for $k(\alpha)$, the cost of monitoring. The adopted regressors are:

[64] For instance, showing that h tries to extract a rent by inflating its income reports would confirm that the inheritance motive matters more than altruism, which cannot be shown definitively by simply looking at the (ambiguous) correlation between h's survey–based income measure and the transfers received.

[65] Frequency of phone contacts is not available for the Mauritanian sample.

[66] We constructed two types of frequency variables: ordinal and increasing, and ratio (number of times per year). The modalities of the former type are produced in Appendix Table 25.

[67] The rationale for considering remittance channels as a proxy for monitoring is that the reception (and, to some extent, use) of in–kind transfers is directly observable by m and involves witnesses (Azam and Gubert, 2004), which in some cases might override the cost of transporting goods oneself to Senegal. Subsequently, one should check for the specific behavior of *in–kind transfers* when dealing with remittances in an imperfect–information world.

dummies for whether h has access to a mobile phone, according to m and to h, and can thus be reached anywhere (*i.e.* these variables *decrease* with communication costs and should therefore enter the regressions *positively*), and a dummy for whether m wrongly stated that h does *not* have a cell phone, which suggests that she always calls the landline, a cheaper option in most cases (this directly proxies a cost and should thus exhibit a *negative* coefficient). Admittedly, these proxies may be correlated with h's income, which in turn affects transfers if altruism or inheritance motives are at work. Tables 18 and 19 therefore control for h's wealth in Columns 5 and 6.

Tables 18 and 19 provide bivariate regressions of total transfers on monitoring variables. Proxies for monitoring activities turn out to be significant, robust in sign and magnitude, and always positive—except, as expected, for the time elapsed since m's last trip to h, since it is a negative function of monitoring.[68] The variables capturing m's contacts with other emigrants from the same household of origin are never significant in the Mauritanian sample. The proxies for $k(\alpha)$ are more volatile, except for m's underestimation of h's access to a mobile phone, which is large in magnitude and significant in the European sample, and exhibits the expected sign.

In spite of endogenity issues, Tables 18 and 19 clearly support the model's prediction of a positive covariance between t and robust, significant proxies for α.[69]

5.2 Misrepresentation at the bottom

The second prediction that we can test empirically is that of larger misrepresentations at the bottom of the distribution of X_{hh}. Note that this does not rely on the assumption that m wants h to increase X_h at least to some threshold X_h^*.[70] The prediction however relies on the assumption, well supported by the data, that X_{hm} and X_{hh} are not independent random variables, *i.e.* that h's misrepresentations are not completely out of touch with reality. Otherwise, a negative correlation between Δ and X_{hh} would

[68] The coefficients are smaller when income or wealth is controlled for, which is evidence of a positive omitted variable bias, but seldom significantly so, which is reassuring. Column 6, which introduces all controls together, often displays insignificant coefficients; this does not come as a surprise given the relatively small sample.

[69] This finding is consistent with De Laat (2008): Migrant husbands in Kenya are "less likely to remit cash if the rural home is far[, *i.e.* if it is not easily monitored, which] suggests efficiency losses".

[70] In effect, if $X_{hh} > X_h^*$, Δ should be negative and increase in absolute value in proportion to how much X_{hh} exceeds X_h^*. So, in both cases (X_{hh} above or below the level preferred by m), Δ and X_{hh} should be negatively correlated.

Table 18: Covariance between monitoring and transfers (European sample)

	(1)	(2)	(3)	(4)	(5)	(6)
Monitoring costs:						
h has access to mobile, *m*'s report	284.8	143.1	317.9	218.2	202.6	73.56
	(279.2)	(276.4)	(279.2)	(270.1)	(281.4)	(269.6)
h has access to mobile, *h*'s report	-46.66	231.4	-136.9	-264.8	-78.35	25.41
	(829.9)	(918.2)	(770.7)	(768.1)	(843.1)	(868.5)
m underestimates *h*'s access to mobile	-1059.4**	-408.6	-1135.7**	-1192.8**	-1030.3**	-542.6
	(460.7)	(508.7)	(448.5)	(515.0)	(467.5)	(559.6)
Monitoring activities:						
Frequency of phone contacts (ordinal; increasing)	633.8***	488.3***	542.0***	495.1***	618.1***	330.1***
	(111.2)	(118.9)	(105.9)	(103.7)	(108.3)	(109.4)
Nb of phone contacts per year	23.90***	19.14***	20.35***	18.61***	23.35***	12.88***
	(4.132)	(4.045)	(4.210)	(4.235)	(4.079)	(4.183)
Frequency of visits to *h* (ordinal; increasing)	343.6***	200.3***	287.8***	266.4***	335.3***	103.3*
	(61.91)	(63.40)	(60.65)	(62.64)	(60.86)	(61.44)
Nb of visits to *h* per year	532.3***	313.1**	456.3***	433.5***	515.1***	193.0
	(136.7)	(136.1)	(134.4)	(139.5)	(134.9)	(134.2)
Time since last visit to *h*	-0.297***	-0.246***	-0.196**	-0.192**	-0.272***	-0.100
	(0.0880)	(0.0897)	(0.0886)	(0.0838)	(0.0886)	(0.0882)
Frequency of contacts with other emigrants from *h* (ordinal)	197.4**	109.2	137.7	180.4**	203.1**	72.82
	(97.49)	(93.45)	(93.09)	(86.94)	(99.10)	(82.20)
Nb contacts with other emigrants from *h* per year	12.57**	8.676*	10.56*	11.75**	12.73**	7.558
	(5.664)	(5.207)	(5.459)	(5.423)	(5.718)	(4.894)
Remit in kind to the household of origin	1262.2***	963.4***	1173.1***	1087.7***	1174.5***	758.2***
	(207.1)	(218.8)	(200.3)	(201.2)	(206.7)	(211.0)

Standard errors in parentheses

The DV is total transfers sent by *m* in the past 12 months (€). Column 1 displays bivariate regressions. Col. 2 controls for *m*'s total labor income; Col. 3 includes a dummy for whether *m*'s spouse lives with *h*; Col. 4 for whether *m* has children in *h*; Col. 5 introduces *h*'s wealth index; and Col. 6 controls for all of the above.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Covariance between monitoring and transfers (Mauritanian sample)

	(1)	(2)	(3)	(4)	(5)	(6)
Monitoring costs:						
h has access to mobile, *m*'s report	76.18	-33.21	42.91	83.93	-21.89	-138.2
	(144.0)	(147.3)	(126.0)	(143.5)	(140.9)	(130.0)
h has access to mobile, *h*'s report	198.1	179.5*	147.0	193.1	132.7	69.57
	(147.9)	(101.1)	(136.5)	(140.3)	(142.8)	(85.31)
m underestimates *h*'s access to mobile	279.6	420.4*	204.1	251.0	246.7	314.3
	(215.9)	(253.5)	(228.5)	(242.1)	(190.9)	(258.6)
Monitoring activities:						
Frequency of visits to *h* (ordinal; increasing)	56.46**	45.01*	41.65*	53.74**	50.75**	28.17
	(24.95)	(24.61)	(24.62)	(25.00)	(23.09)	(23.01)
Nb of visits to *h* per year	29.94	26.86	14.60	25.57	28.04	12.71
	(21.29)	(19.83)	(20.92)	(21.65)	(20.09)	(18.73)
Time since last visit to *h*	-0.136***	-0.119**	-0.0816	-0.120**	-0.134***	-0.0764
	(0.0523)	(0.0481)	(0.0497)	(0.0510)	(0.0516)	(0.0467)
Frequency of contacts with other emigrants from *h* (ordinal)	66.44	46.30	72.98	78.30	63.38	39.80
	(58.02)	(46.85)	(63.98)	(67.46)	(67.02)	(58.30)
Nb contacts with other emigrants from *h* per year	5.823	4.225	5.672	6.670	6.126	3.430
	(6.103)	(4.603)	(6.938)	(7.076)	(6.781)	(5.802)
Remit in kind to the household of origin	338.8***	252.9***	289.5***	294.6***	310.9***	221.7**
	(98.66)	(89.33)	(95.59)	(101.5)	(97.94)	(87.25)

Standard errors in parentheses

The DV is total transfers sent by *m* in the past 12 months (€). Column 1 displays bivariate regressions. Col. 2 controls for *m*'s total labor income; Col. 3 includes a dummy for whether *m*'s spouse lives with *h*; Col. 4 for whether *m* has children in *h*; Col. 5 introduces *h*'s wealth index; and Col. 6 controls for all of the above.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

obtain trivially.[71]

Figures 2 and 3 present a selection of asset ownership and housing quality variables.[72] The figures graph Δ against X_{hh} and all show a sharp negative relationship, supporting the prediction of larger discrepancies at the bottom of the X distribution, or more accurately: the more serious the infringement of m's earmarking instructions. This result is robust to using a quadratic fit and Lowess estimation, as well as to the removal of outliers (not reported).

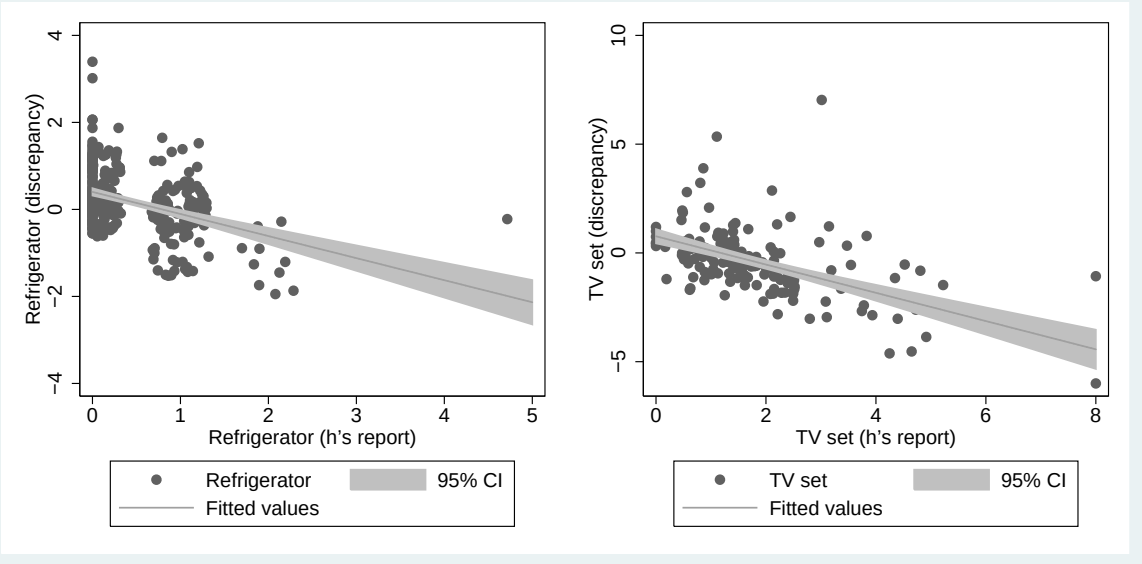

Figure 2: Relationship between Δ and X_{hh}: Refrigerator and TV set ownership

5.3 Relationship between transfers and misrepresentations

Indirect empirical evidence Finally, the model implies that rent extraction could be further established through the comparison of the signs of Δ and $t'(\cdot)$. This is however fraught with difficulties. First, a regression of t on X_{hm} would be a challenge to interpret since one key conclusion of our discussion is that X_{hm} is a *decision* variable for transfer–maximizing h. Thus, it is endogenous to t, as well as all the other variables in the model (α, a and X_{hh}). In the absence of a passable instrument, we cannot

[71] Misreporting at the bottom, *i.e.* a negative correlation between Δ and X_{hh} obtains whenever

$$X_{hm} = a + bX_{hh} + \epsilon, \quad \text{and} \quad -1 < b < 1, \quad b \neq 0.$$

[72] The interest reader is invited to consult Seror (2012) for further examples.

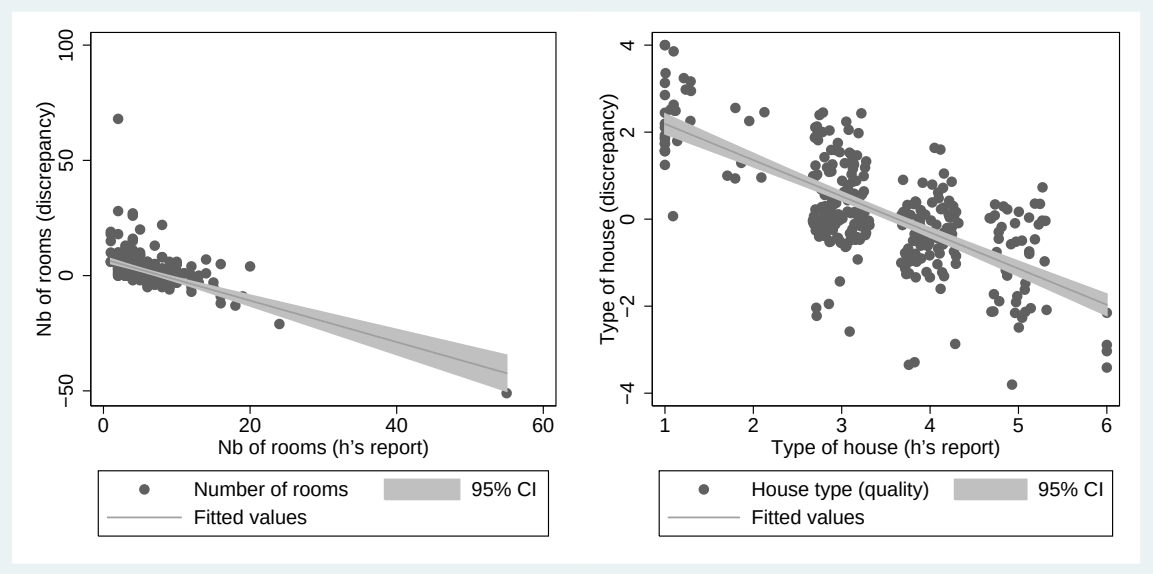

Figure 3: Relationship between Δ and X_{hh}: Number of rooms in h's dwelling and House type (quality)

causally identify the sign of $t'(\cdot)$. We shall therefore only rely on indirect or non–econometric evidence.[73]

Second, the data provide no clear–cut information on whether transfers are earmarked for particular expenditures. We have thus far taken all X_h variables for which we have matched data as behaving more or less in the same way. However, it may well be that, for instance, migrants think (on average) their households of origin have enough TV sets and too few refrigerators, cars, etc. In this case, we should observe a Δ*TV set* and a $t'(\cdot)$ of the same (negative) sign. This is indeed what might be at work: TV sets in our data behave as other assets (especially, the "misreporting at the bottom" prediction is fulfilled), but Δ*TV set* is consistently negative. The upshot is that the coefficient on TV set in a transfer regression would pick up the effect of the *other* items that m paid for but h sold or did not buy.[74]

An indirect piece of evidence comes from Section 3.3. Indeed, the fact that the discrepancy in the age of h's head is explained by m's being mistaken about his or her identity and explains other discrepancies, is suggestive evidence of rent extraction

[73] Simple, regression–based correlations of t and X_{hm} are provided in Appendix Tables 26 and 27.

[74] The coefficient would thus be positively biased as ownership of assets already owned is likely to be highly positively correlated.

insofar as this is the most likely explanation: Younger household members are in charge of h and secure transfers from m by misrepresenting the actual chain of command; this enables them to do with t whatever they think best without arousing m's suspicion.

Semi–qualitative evidence The follow–up survey carried out in October 2012 asked respondents to speculate on the results emerging from the MIDDAS and migration–PSF data. 17 out of 20 interviewees maintained that at least one of the discrepancy categories (concerning asset ownership, housing characteristics or the number of emigrants remitting to h) are due to h lying to m in order to extract more or secure transfers. This result is particularly strong for assets: 13 out of 20 respondents blame the Δ's on h's disingenuousness.[75] The most frequent story pertains to durable goods not being purchased, contrary to m's wishes, or sold, if sent in–kind. The respondents often mention the purchase of fancy clothes or participation in "baby naming ceremonies" (*ngente* in Wolof) as the destination of the embezzled transfers.[76]

As one respondent put it: "We only know what they tell us". Although the interviewees did not always link phone contacts with monitoring, half the sample has doubts about the information they receive from their main transfer recipients, and think that their earmarking is not followed or that information is distorted in order to extract rents. Accordingly, the amounts spent on calling h are sometimes exorbitant. The mean per month is €26.44 for those without a spouse or child living with h. But several cited figures above €40 or even in the hundreds. The average is expectedly much higher for those with a spouse or child in h: above €70 per month, which may reflect sentimental ties or the fact that migrants care more about what happens (and how the money is used) when they transfer to a spouse or child.[77]

As mentioned in Section 4.2, although only one respondent acknowledged having recently cut transfers and badmouthed the recipient in retaliation for information dis-

[75] 7 of those understood the question immediately and spontaneously answered without the enumerator's listing possible options; 12 came up with an anecdote or an example to illustrate their answers.

[76] Although ideally we would like to hear h's hypotheses about the discrepancies, it is worth noting that migrants do not systematically accuse h. Far from it: 7 respondents said (among their three most plausible explanations, for at least one discrepancy category) that m would exaggerate the impact of her transfers, and 5 that she would insist on her coming from a well–off background, and thus inflated X_{hm}.

[77] These amounts are consistent with De Laat's (2008) conclusion on Kenyan split couples: "Split migrant men invest substantial resources and time monitoring their rural wives by making (costly) short, frequent visits home". However, De Laat's study focusing on split couples, he cannot disentangle the sentimental and monitoring purposes of visits.

tortion and misuse of his transfers, 6 out of 10 respondents said they would or could cut transfers if faced with misrepresentations.

Socio–anthropological evidence The anthropological literature also supports the rent–extraction hypothesis. In effect, there seems to be a strong social pressure on Senegalese migrants to remit: "Whereas a lack of economic success is not deemed unusual in Africa, for a migrant in Europe it is perceived as failure and brings with it social contempt" (Gerdes, 2007), which encourages them to embellish their life stories, and allows a potentially profiteering attitude from transfer recipients.

Besides, migrants often want to invest in their home country. But lack of trust in one's business partners back home seems pervasive among Senegalese migrants: They indeed "quite often bemoan the fact that no one is trustworthy in their family circles and friends" (Fall, 2003, translation by the author).

The importance of monitoring and networks for expanding migrants' information set further supports the hypothesis that rent extraction is an issue. As Dia (2009, cited in Chort et al. (2012)) puts forward, both m and h are controlled through reputation and the use of rumors in "a permanent adjusting or re–adjusting mechanism for individual behaviors within the group" (Chort et al., 2012) that is only necessary if behavioral rules (remitting, truthfully representing one's situation to the remittance sender, etc.) are indeed sometimes infringed.[78]

6 Sources of bias in transfer regressions

6.1 Absence of bias under perfect information

Sections 3 and 5 were meant to show empirically that information asymmetries and rent extraction can be a concern within multi–sited households. We shall now investigate what this entails for empirical and theoretical works on migrants' remittances.

Most of the empirical literature on remittances resorts to unmatched data: Either m's or h's reports are known, for both m's and h's characteristics. Osili (2007) uses a matched dataset, and combines information from m and h in one transfer equation. Her

[78] Azam and Gubert (2004) report that rent extraction by h is not only an isolated case but can also give rise to peer pressure to exploit the informational rent and shirk when the lie requires collusion: "Acts of collusion, such as announcing fictitious natural disasters, were actually observed in the Kayes region" (*ibid.*).

rationale is that matched data can help us avoid measurement error: In particular, "self–reported measures of origin household assets [...] may represent less noisy estimates of the origin household's economic position".

This approach is valid as long as one assumes away information asymmetries through misrepresentations. Then X_{hm} and X_{hh} are equivalent except that the latter should suffer less from measurement error. Conversely, if h is able to manipulate information, there is no reason to expect X_{hm} and X_{hh} to yield similar results. Now, neither X_{hm} nor X_{hh} can be construed as the "legitimate" X_h:[79] X_{hh} may be argued the true value of X_h; but since X_{hm} is likely to be manipulated (it is h's choice variable), X_{hh} has an impact on t only insofar as it is not totally disconnected from X_{hm} (which is what m believes and bases her remittance decisions on). Therefore, if information can be manipulated, transfers regressions are likely fraught with simultaneity bias when X_{hm} is used, and both omitted variable bias and reverse causation with X_{hh} on the right–hand side.

Intuitively, the bias will depend on the sign of the relationship between t and X_h. Even if we were concerned with remittance motives, economic theory might not be able to tell us the sign of $t'(X_h)$, as it is not always clear that one particular variable X_h relates to a single motive.[80] In our earmarking framework, $t'(X_h)$ depends on whether m wants X_h to increase: In this case, m sends transfers, *e.g.* for building, repairing or purchasing, thus increasing the quality of the housing (be it h's own property or m's, which h can use) or the number of assets owned by h.

This is why in what follows we consider only the case where $t'(X_h) > 0$ and $\Delta > 0$, *i.e.* rent extraction with m willing to increase X_h. We further assume for tractability m's preferences (the level of X_h she favors for h) to be exogenous.

[79] This stands in stark contrast to McKenzie et al. (2006) and LaLonde (1986), who compare experimental, "true" estimates, with non–experimental, *a priori* biased ones.

[80] For instance, the quality of h's dwelling can reflect m's investment in that particular aspect of their wellbeing or her investment in real estate ($t'(X_h) > 0$), or it can reflect h's living standard, which induces lower (higher) remittances if m is altruistic (invests in her return or inheritance).

6.2 Derivation of the bias

More formally, regressing t on X_{hm} amounts to estimating Equation 4a from the following system, ignoring that X_{hm} is in fact a choice variable for h:

$$t = \beta_{11} X_{hm} + u_1 \tag{4a}$$

$$X_{hm} = \beta_{21} t + \beta_{22} p(\Delta;\alpha) + \beta_{23} X_{hh} + u_2 \tag{4b}$$

$$X_{hh} = \beta_{31} t + \beta_{32} X_{hh_0} + u_3 \tag{4c}$$

$$p(\Delta;\alpha) = \beta_{41}\Delta + \beta_{42} k(\alpha) + u_4 \tag{4d}$$

The notation is the same as in the theoretical model. $p(\Delta;\alpha)$ is a positive function of misrepresentation and investment in monitoring, α, which comes at an exogenous cost $k(\alpha)$. Finally, it is likely that transfers influence the current level of X_{hh} but this remains partially determined by previous levels of X_{hh}—X_{hh_0} stands for the initial X_{hh}, before remittances started. Note that X_{hh} does not appear in Equation 4a because we assumed that transfers are 0 if fraud is detected, as in the theoretical model.

Let us abstract from the issue of the impact of past transfers on X_{hh} (Equation 4c) and derive the bias in β_{11} when Equation 4a is estimated on its own, *i.e.* when information asymmetries are assumed away.

Substituting for t and $p(\Delta;\alpha)$ into Equation 4b:

$$X_{hm} = \beta_{21}(\beta_{11} X_{hm} + u_1) + \beta_{22}(\beta_{41}\Delta + \beta_{42} k(\alpha) + u_4) + \beta_{23} X_{hh} + u_2 \tag{5a}$$

$$X_{hm} = (\beta_{21}\beta_{11} + \beta_{22}\beta_{41}) X_{hm} + \beta_{22}\beta_{42} k(\alpha) + (\beta_{23} - \beta_{22}\beta_{41}) X_{hh} + \beta_{21} u_1 + u_2 + \beta_{22} u_4 \tag{5b}$$

$$(1 - \beta_{21}\beta_{11} - \beta_{22}\beta_{41}) X_{hm} = \beta_{22}\beta_{42} k(\alpha) + (\beta_{23} - \beta_{22}\beta_{41}) X_{hh} + \beta_{21} u_1 + u_2 + \beta_{22} u_4 \tag{5c}$$

Assuming $\beta_{21}\beta_{11} + \beta_{22}\beta_{41} \neq 1$[81]:

$$X_{hm} = \frac{\beta_{22}\beta_{42}}{1 - \beta_{21}\beta_{11} - \beta_{22}\beta_{41}} k(\alpha) + \frac{\beta_{23} - \beta_{22}\beta_{41}}{1 - \beta_{21}\beta_{11} - \beta_{22}\beta_{41}} X_{hh} + \frac{\beta_{21} u_1 + u_2 + \beta_{22} u_4}{1 - \beta_{21}\beta_{11} - \beta_{22}\beta_{41}} \tag{6}$$

For the sake of clarity, let us assume zero correlation between u_1, u_2 and u_4. Then

[81] Under this assumption, reduced forms for X_{hm} and t exist.

we have:

$$Cov(X_{hm}, u_1) = \frac{\beta_{21}}{1 - \beta_{21}\beta_{11} - \beta_{22}\beta_{41}} \mathbf{E}u_1^2 \tag{7}$$

As can be seen from Equation 7, the sign of the (simultaneity) bias from estimating Equation 4a alone is indeterminate, but there is no reason to assume it to be nil. When $t'(X_h) > 0$, β_{11} and β_{21} are positive. Besides, we expect β_{22} in Equation 4b to be negative as an increased probability of detection should reduce h's incentive to inflate X_{hm}. Conversely, β_{41} is assumed to be positive in the model because blatant misrepresentations are easier to uncover. The sign of $1 - \beta_{21}\beta_{11} - \beta_{22}\beta_{41}$ cannot be straightforwardly derived, as we have no way of knowing *a priori* the magnitude of the β parameters.

What happens when Equation 4a is estimated separately and X_{hh} is used instead of X_{hm}? Replacing X_{hm} by X_{hh} and estimating $\widetilde{\beta_{11}}$, the sign of the inconsistency in $\widetilde{\beta_{11}}$ is given by:

$$- \frac{\beta_{21}}{\beta_{21}\beta_{11} + \beta_{23} - \beta_{22}\beta_{41}} \tag{8}$$

Contrary to what happened with X_{hm} in Equation 4a, the sign of the bias is now unambiguously negative. But interestingly, the bias need not be of opposite signs in Equations 7 and 8. Indeed, if $\beta_{21}\beta_{11} + \beta_{22}\beta_{41} > 1$, the former is negative too. However, if $\beta_{21}\beta_{11} + \beta_{22}\beta_{41} > 1$ we can say that the bias in Equation 7 is more negative than that in 8 because

$$\frac{\beta_{21}}{1 - \beta_{21}\beta_{11} - \beta_{22}\beta_{41}} < -\frac{\beta_{21}}{\beta_{21}\beta_{11} + \beta_{23} - \beta_{22}\beta_{41}} < 0 \tag{9}$$

holds if and only if

$$1 + \beta_{23} > 2\beta_{22}\beta_{41} \tag{10}$$

which is always true since $\beta_{22} < 0$.

To conclude on the bias derivation, transfer equations that rely on the significance of a variable to determine remittance motives may yield very different results depending on whether they regress transfers on X_{hm} or X_{hh}. But contrary to X_{hm}, the bias in X_{hh} has the property of being always non–positive, which can be useful when the

alternative hypothesis is that the coefficient on X_h is positive. For instance, if one intends to test the inheritance motive of remittances (Hoddinott, 1994, *inter alia*), *i.e.* that h's holdings enter transfer regressions positively, using h's instead of m's report has the advantage of offering a lower bound interpretation.

6.3 Comparison of transfer equations using migrants' and households' reports

Comparing transfer equations using migrants' and households' reports can only be illustrative. In effect, several elements preclude diagnosing the observed differences in terms of simultaneity and omitted–variable biases. First, because we abstracted from the measurement error between X_{hm} and X_{hh}, which should entail an attenuation bias in the former. Our small sample aggravates this issue. Second, because of the ambiguity surrounding the direction of the bias, even though necessary simplifying assumptions were introduced. Third, because the direction of the bias depends on the motive for misrepresentations, and in our earmarking context on whether on average X_{hh} is in excess or falls short of m's instructions.

Figure 4 reproduces for the record Osili's (2007) transfer equations. Note that she does *not* compare m's and h's reports of h's "asset variables"[82] explicitly, since she uses the former only when the dependent variable is the remittances *sent* by m whereas the latter corresponds to "remittances received in the past year by the origin family member" on the left–hand side.[83] Besides, Columns 1 and 2 use both "asset variables" at the same time, contrary to Columns 3 and 4.

Accepting the results as "comparable" nevertheless, we see from Figure 4 that Osili's model–based hypothesis of a negative correlation between h's "asset variables" and transfers cannot be unambiguously shown in this specification. Indeed, the significant, negative coefficient on landholdings (Col. 4) could be due to a negative bias—which is all the more plausible as X_{hh} is used. A bias due to ignored information asymmetry could also be at work for the number of buildings: If the true coefficient is 0, the point estimate could still be negative (Col. 3) and even significantly so (Col.

[82] Namely, "the size of the origin household's land holdings (measured in ha) and the number of buildings owned by the origin family".

[83] In Osili's (2007) own words, "because migrants send transfers to a complex web of family members in the origin country, estimates based on transfers sent by the migrant and transfers received by a given origin family are comparable but may not be identical".

Table 3
Transfers to the origin family

Dependent variable	U.S. migrants	U.S. migrants and origin households		
	Transfer sent to home family ($)	Transfer sent to home family ($)	Transfer received ($) by origin family	Transfer received ($) by origin family
	(1)	(2)	(3)	(4)
	Migrant reports only		Migrant and origin family reports	
Migrant variables				
Migrant's age (at the time of the survey)	−25.14 (64.03)	−59.78 (98.23)	14.68 (106.22)	−3.38 (162.22)
No. of children in household	−487.14 * (281.49)	−886.21 * (490.32)	−839.62 **	−1038.02 * (559.35)
Years of schooling	39.41 (250.01)	432.89 (560.91)	632.33 403.916	867.40 556.4596
Migrant's occupation (skilled = 1)	1984.26 ** (883.31)	2173.23 (1187.24)	958.22 (1161.23)	1345.42 (1391.05)
Household income ($\times 10^3$)	0.02 * (0.01)	0.01 (0.01)	0.02 ** (0.01)	0.02 * (0.01)
Origin family variables				
(Migrant report)				
Origin family size	272.06 *** (75.33)	649.93 *** (247.79)	293.18 (231.22)	347.05 (342.29)
Rural	−699.78 (893.54)	−949.05 (1525.46)	−1457.97 (1391.86)	−1338.50 (1705.26)
Landholdings (ha)	−68.61 (46.00)	−68.44 (307.74)		
No. of buildings	−188.59 *** (59.34)	−137.41 (144.08)		
(Origin family reports)				
Landholdings (ha)				−1013.85 ** (455.22)
No. of buildings			−118.34 (330.49)	
Age of the head			39.20 (46.76)	84.47 (68.32)
Migrant is child of respondent = 1			1310.16 (996.57)	934.78 (1481.23)
Constant	37.32 (4652.70)	−7988.78 (9718.84)	−13001.20 * (7541.10)	−18489.58 * (10473.47)
No. of observations	100	57	55	41
Adjusted R^2	0.27	0.43	0.48	0.57
Sample	Migrant	Matched	Matched	Matched

Robust standard errors are shown in parentheses.
Transfers sent refer to total remittances sent by a migrant to ALL family members in the origin country and is based on the migrant's report.
Transfers received by the origin family refer to remittances received by a specific origin family member from a given U.S. migrant and is based on the origin family's report.
* Represents 10% level of significance.
** Represents 5% level of significance.
*** Represents 1% level of significance.

Figure 4: Osili's (2007) transfer equations

1) when X_{hm} is used, provided the bias goes in the same direction—and is therefore stronger.

Tables 20 and 21 (for the European and Mauritanian samples, respectively) both contain three columns. The first one presents the regression of total transfers on migrant and household characteristics for the whole (matched and unmatched) sample.[84] Col. 1 thus uses the values reported by the migrant. Col. 2 focuses on the matched sample, while still using migrants' reports. Finally, Col. 3 displays the regression run on the same sample as Col. 2 but replaces household–characteristic variables with h's reports.

The variables on the right–hand side were chosen so as to follow the regressions in Figure 4 as closely as possible.

Some of the coefficients in the top half of Tables 20 and 21 are significant and stable (*e.g.* age and income in the European sample, and the number of m's children living with h in the Mauritanian one, respectively), whereas others are more volatile. Since Col. 1 and 2 use exactly the same variables and only differ by the samples on which the regressions are run, there are two possible explanations: selection into the matched sample and measurement error. As we saw in Table 4 that the former is not very serious, we interpret this as evidence of the latter.

As far as the bottom half is concerned, where household characteristics—size of h, age of h's head and house type—are measured using the migrant's (their own) reports in Col. 2 (3), coefficients are never significant beyond the first column in Table 20, hinting at serious measurement error. An endogeneity bias of very different strengths (depending on whether X_{hm} or X_{hh} is used) might be at work in Table 21 where the age of h's head is alternately positive and significant or negative and insignificant. Unfortunately, it is not possible to disentangle empirically what is to be blamed on measurement error and on endogeneity bias. If the simplifying assumptions in Section 6.2 are legitimate, Table 21 suggests that the age of h's head is a significant predictor of transfers as the bias in X_{hh} could not be positive; the bias in X_{hm} would also be negative but much stronger in magnitude.

[84]The size of the unmatched sample is severely reduced by missing values. This is especially true of the European sample, where remittances and income are often missing. However, the results are robust to omitting the latter and thus increasing the sample by 152 observations (not reported).

Table 20: Transfer equations based on Osili (2007), European sample

	(1) All, *m*'s report	(2) Matched, *m*'s report	(3) Matched, *h*'s report
Migrant's characteristics (*m*'s report):			
Age	46.94***	44.01**	45.05**
	(9.973)	(19.16)	(18.09)
Number of resident children	-224.7**	0.286	-270.9*
	(89.47)	(241.7)	(138.3)
Schooling	-47.30	-38.79	9.901
	(52.27)	(110.1)	(110.9)
Total household income (per month, €1000)	547.0***	736.6***	666.4***
	(114.1)	(218.3)	(199.2)
Son/Daughter of *h*'s head (d)	516.3***	243.3	579.6*
	(191.6)	(424.9)	(330.4)
Household's characteristics (*m*'s report):			
Rural (d)	535.1**	901.1*	524.2
	(271.2)	(516.0)	(436.6)
Household's characteristics (*m*'s or *h*'s):			
Size of household of origin	4.635	-15.63	-5.620
	(10.38)	(21.25)	(19.15)
House type (quality)	168.3**	102.0	9.477
	(83.65)	(207.3)	(128.8)
Age of *h*'s head	-15.35*	-0.633	-17.32
	(8.285)	(17.60)	(12.85)
Constant	-561.8	-1021.5	8.473
	(689.9)	(1443.6)	(1145.7)
Observations	372	118	125

Standard errors in parentheses
The dependent variable is the total amount of transfers sent by *m* to *h* during the year before the MIDDAS survey. Schooling ranges from 0 for no schooling to 5 for college education.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 21: Transfer equations based on Osili (2007), Mauritanian sample

	(1) All, *m*'s report	(2) Matched, *m*'s report	(3) Matched, *h*'s report
Migrant's characteristics (*m*'s report):			
Age	14.83***	6.105	-4.509
	(5.697)	(7.897)	(6.617)
Number of resident children	-87.64***	-79.74**	-87.21**
	(24.37)	(34.69)	(37.00)
Schooling	1.462	26.64	4.230
	(43.58)	(64.31)	(57.02)
Total household income (per month, €1000)	1189.8**	811.5	753.5
	(557.0)	(527.3)	(520.3)
Son/Daughter of *h*'s head (d)	240.5**	80.00	-257.5
	(105.8)	(144.1)	(189.2)
Household's characteristics (*m*'s report):			
Rural (d)	-33.27	-88.11	-77.82
	(101.9)	(151.9)	(150.4)
Household's characteristics (*m*'s or *h*'s):			
Size of household of origin	3.468	1.468	-4.355
	(6.419)	(10.28)	(10.00)
House type (quality)	-21.61	26.18	38.62
	(66.31)	(98.36)	(59.84)
Age of *h*'s head	-8.509	-9.686	16.75*
	(5.747)	(10.22)	(8.693)
Constant	260.5	716.7	-200.8
	(301.7)	(507.9)	(528.9)
Observations	245	141	137

Standard errors in parentheses
The dependent variable is the total amount of transfers sent by *m* to *h* during the year before the MIDDAS survey. Schooling ranges from 0 for no schooling to 5 for college education.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Measurement error can also be seen from the standard errors, especially in Table 20 where they are almost twice as high in the matched sample, compared to Col. 1 where both the unmatched and matched observations are used. This is due to the fact that sample size in Col. 2 and 3 is slightly more than 30% of that in Col. 1—as against more than 55% for the Mauritanian sample.

Except for sample size issues and a selection bias unlikely to be serious, the results in all three columns should be roughly similar in the absence of information asymmetry. Col. 2 and 3 should exhibit the same coefficients on X_h, save for higher measurement error in the former.

In order to show more firmly that the bias in X_{hm} and X_{hh} need not be of the same sign or strength, Tables 22 and 23 display simple bivariate[85] regression coefficients. The dependent variable is total transfers and the regressors are variables for which we have both m's and h's reports.

The coefficients are quite sensitive to the use of X_{hm} or X_{hh}, changing signs and jumping in and out of confidence intervals, but they are seldom statistically significant.

7 Discussion and recommendations

7.1 Taking information asymmetry into account in transfer regressions

In order to close the rent–extraction argument, and also because most of the economic literature on remittances focuses on the sign and significance of various coefficients to test competing theories, we would ideally estimate $t'(\cdot)$ empirically.[86] Although we do not have an exogenous instrument for X_{hm} or X_{hh}, necessary to carry out this exercise, we now present potential identification strategies that enabled us to estimate $t'(\cdot)$:

First, surveys need to be more precise about the remittance decision process: What does m base her transfer decisions on? Does she (explicitly or implicitly) impose a consumption pattern on h? Does she monitor or punish h in case this earmarking is not followed? To what extent are the regressors (say, particular assets) proxies for other,

[85] Although uncentered variance inflation factors show low multicollinearity (not reported), the matched variables considered might be correlated with each other, thus introducing various endogeneity issues in a multivariate regression.

[86] We saw that information asymmetry must be taken into account, and Δ and $t'(X_h)$ of the same sign would help prove that h "taxes" X_h in favor of another consumption item.

Table 22: Bivariate transfer equations with matched asset variables

	(1) Migrant's report	(2) Household's report
fridge	-123.1	-31.88
	(129.1)	(151.3)
freezer	176.6	56.42
	(172.4)	(253.2)
tvset	110.4	42.31
	(83.72)	(69.93)
dvd	44.77	196.3
	(112.6)	(182.8)
radio	146.8**	67.78
	(68.07)	(60.56)
cd	116.3	-9.444
	(114.4)	(95.61)
fan	30.84	141.6**
	(58.17)	(65.01)
plow	331.2**	-119.7
	(166.7)	(194.6)
seeddrill	121.7	-280.7
	(248.4)	(239.1)
hoe	110.6	18.51
	(69.54)	(37.96)
car	64.82	30.77
	(174.5)	(292.1)
mopedscooter	-53.83	-64.57
	(178.1)	(359.9)
motorbike	220.9	-465.4*
	(280.1)	(262.9)
bike	144.4	-82.75
	(135.7)	(119.6)
cart	68.00	0.212
	(123.3)	(235.9)
canoe	483.1	41.15
	(510.3)	(217.6)

Standard errors in parentheses

A Mauritanian dummy is included in each regression to allow for a lower remitting capacity (coefficient not reported).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23: Bivariate transfer equations with matched housing variables

	(1) Migrant's report	(2) Household's report
House type (quality)	40.50	-56.72
	(103.0)	(83.88)
Number of rooms	50.60***	-0.238
	(15.23)	(21.78)
Roof material (quality)	109.9	-0.696
	(104.6)	(107.9)
h owns their dwelling	37.75	-111.2
	(442.7)	(262.8)

Standard errors in parentheses

A Mauritanian dummy is included in each regression to allow for a lower remitting capacity (coefficient not reported).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

more aggregate variables (living standards, for instance) that are less straightforwardly quantified? Ideally, migrants could be surveyed over an extended period of time and asked to keep an accurate account of the monies they send and what they are meant to finance.

Second, measuring exogenous variation in monitoring costs can help identify X_{hm}. Mobile network coverage and quality, the exogenous increase in the tax on incoming calls in Senegal between August 24, 2011 and May 15, 2012,[87] etc., are possible instruments for X_{hm}, as they affect it (through the frequency of mobile phone contacts) but do not enter the transfer decision directly, since they are due to exogenous variation (geography and presidential decrees, respectively) in h's environment. It is also possible to randomize such monitoring costs. This approach has been developing recently (Ashraf et al., 2011, for instance). But the limitation of this identification strategy is that one still needs to gather data on discrepancies between what m believes and the actual state of affairs in h. Otherwise, it is difficult to argue convincingly that the (causal) impact of communication prices on transfers indeed occurs through monitoring and misrepresentations.

[87] Unfortunately, both MIDDAS and migration–PSF data were collected before the tax was introduced. We were not able to access network coverage data.

An alternative to collecting costly matched data, while still making sure that information asymmetry is indeed at work, is a to run a detailed survey on the changes in the migrant's behavior induced by the shift in monitoring costs. For instance, if phone credits are randomly distributed to a treatment group, treated migrants might start calling relatives they previously had little contact with. They might also be updated about the recipients' financial needs rather than using the extra time they can afford to monitor them. Such experiments must thus plan survey questions on *whom* the treated migrants called and *what* they talked about, and be able to compare pre– and post–treatment behaviors.[88]

7.2 Two–way information asymmetry

A natural extension of the model and empirics is to take account of misrepresentations by m to h rather than abstract from them and assume that information asymmetries only occur in the opposite direction. Unfortunately, the structure of the data does not enable us to test for information asymmetries created by the migrant, although "[insecure migrants] are found to remit at least as much and very likely more than migrants with less precarious working conditions" (Chort et al., 2012), which suggests a social pressure to remit and be successful, and thus an incentive to give a favorable account of one's economic situation as a migrant. In this case, m's misrepresentations would be "altruistic" or meant to save face. Indeed, 14 out of the 20 respondents to the semi–qualitative survey mentioned having hidden or distorted information to h; 5 of them had done it at least once in the past twelve months. The main reason behind this is that m does not want h to be worried about her. A secondary explanation is that m wants to "save face" and will not admit to having health or economic problems.

Conversely, "escape clauses" (Azam and Gubert, 2004) might allow m to alleviate her burden and exaggerate the frequency of illnesses and unemployment spells so as to extract a rent and escape her obligations, after having received financial support from h to emigrate and settle in a foreign country. Only one interviewee acknowledged

[88] Besides, randomized field experiments need not confine themselves to communication costs. Giving to a treatment group information about the Senegalese associations present near their places of abode or paying their subscription fees to encourage them to register can help exogenously expand migrants' information set about recipients' actions, and therefore affect their transfer behaviors. However, detailed data must again be collected about the nature of their interactions with their fellow association members.

having distorted information[89] in order to free himself temporarily from the obligation to remit. But another put forward that friends of his do it but will not admit to it as it is "a serious matter". Non–altruistic misrepresentations from m might thus be more common than what the survey bespeaks because it is considered morally reprehensible.

However, the social pressure seems strong enough that nearly 80% of migrants in the matched sample remit regularly. In the semi–qualitative data, all respondents remit, 15 of them do so on a regular basis and 12 feel remitting is a "moral obligation".

We see the collection of matched data on m's misrepresentations as an important complement to our analysis of the information asymmetries created by h. m might be able to become more successful in the host country, and thus remit more, by resisting part of h's requests and invest in a small business, a car (or simply a driver's license) or any other action that would enhance her job prospects. If social pressure to remit prevails, transfers would be suboptimally high for m and a windfall for h. Indeed, we have some evidence of this in the semi–qualitative data, as 8 respondents deem the remittances they send a "hindrance to success in France". In the opposite case ("escape clauses"), it is unclear *a priori* which, of m and h's misrepresentations, should dominate—we expect the equilibrium to depend on m and h's respective monitoring capacities and bargaining powers—and it is not clear either whose misrepresentations are less detrimental to the well–being of the multi–located household.

In effect, it does not follow from the evidence on rent extraction that transnational households would be better off if information asymmetries could be avoided. Although rent extraction leads to "efficiency losses from costly information acquisition" (De Laat, 2008) and migrants tend to denounce the frivolous use made by h of diverted remittances, it might well be that m tries to impose an inadequate consumption pattern on h: "Those new figures of success array themselves in new and incontestably conspicuous material attributes: handsome mansions, parabolic antennae, luxury cars, electrical domestic appliances", etc. (Dia, 2007, translation by the author).[90] Migrants

[89] Namely, he exaggerated the time he spent "without a roof over [his] head".

[90] Dia (2007) further notes: "It is fashionable among migrants nowadays to have a house built in the village [even though the existing house might be large enough]. Oftimes, building does not suffice; the house must be adorned with all the attributes of 'modernity': a TV set, a VCR, a telephone, and electrification thanks to solar pannels."

are therefore often "accused of imposing decisions on the villagers and thus of abusing their new monetary power" (*ibid.*).

8 Conclusion

Matched data on Senegalese migrants in France, Italy and Mauritania enable us to contrast migrants' descriptions of the characteristics of their households of origin with what those most directly concerned themselves report. We can thus compute the discrepancies between migrants' and households' responses to identical survey questions. Assessing and rejecting alternative hypotheses, we conclude that the discrepancies are systematically non–zero and most probably due to information asymmetry, and more precisely (in the light of empirical evidence and the socio–anthropological literature on Senegalese migrations): rent extraction by households of origin. We then develop a soft–information model tailored to the migration context. The model is invoked to provide empirical predictions. Most of them can be tested and support the information–asymmetry interpretation of the observed discrepancies. The basic assumptions of the model are used to simplify computations and derive the bias induced in transfer regressions when information asymmetry is ignored. Whereas the empirical literature on migrants' remittances systematically assumes away information asymmetry in migrant–household pairs, and investigates the determinants of transfers on the basis of the sign and significance of regressors, we show that using migrants' or households' reports need not lead to the same results, both yielding severely biased estimates but potentially pointing in opposite directions.

Showing that information asymmetries are pervasive between migrants and households has at least two consequences. First, it is further evidence of the empirical irrelevance of the unitary and collective household frameworks. In effect, insofar as migrants and their relatives in the home country are regarded as members of the same (transnational) household, their decisions should benefit the group as a whole. However, since households of origin seem able and willing to extract rents from remittance senders thanks to private information they can manipulate, the assumption or prediction (based on altruism) that migrants and their non–migrant relatives should necessarily arrive at

efficient intra–household allocations is unwarranted.

But one should not be too hasty in drawing policy recommendations. Indeed, and quite paradoxically, as implied by the model, reducing information asymmetry between migrants and their households of origin might *curtail* remittances. Estimating this prediction causally, for instance thanks to randomized field experiments, while making sure that transfer behavior is altered through misrepresentations and monitoring, is a promising avenue for future research. What should resolve the question whether information asymmetry in the remittance framework is detrimental to development is the impact on development of transfers used according to the migrant's wishes *versus* following the household's assessment of their own needs.

Bibliography

Ashraf, N., D. Aycinena, and D. Martinez, C. and (2011). Remittances and the problem of control: A field experiment among migrants from El Salvador. *Working Papers*.

Azam, J. and F. Gubert (2004). Those in Kayes: The impact of remittances on their recipients in Africa. *IDEI Working Paper 308*.

Brown, S., S. Hillegeist, and K. Lo (2004). Conference calls and information asymmetry. *Journal of Accounting and Economics 37*(3), 343–366.

Carling, J. (2008). The determinants of migrant remittances. *Oxford Review of Economic Policy 24*(3), 581–598.

Chort, I., F. Gubert, and J.-N. Senne (2012). Migrant networks as a basis for social control: Remittance incentives among Senegalese in France and Italy. *Regional Science and Urban Economics*.

De la Brière, B., E. Sadoulet, A. De Janvry, and S. Lambert (2002). The roles of destination, gender, and household composition in explaining remittances: an analysis for the dominican sierra. *Journal of Development Economics 68*(2), 309–328.

De Laat, J. (2008). Household allocations and endogenous information. *CIRPEE Working Paper No. 08-27*.

De Vreyer, P., S. Lambert, and A. Safir (2008). Remittances and Poverty: Who benefits in the Household? *Paris School of Economics mimeograph*.

Dia, H. (2007). Les investissements des migrants dans la vallée du fleuve Sénégal: confiance et conflits d'intérêts. *Revue européenne des migrations internationales 23*(3), 29–49.

Fall, P. (2003). Migration internationale et droits des travailleurs au Sénégal. *Série UNESCO: Rapports par pays sur la ratification de la Convention des Nations Unies sur les droits des migrants*.

Gerdes, F. (2007). Focus migration. Senegal. *Hamburg Institute of International Economics 10*.

Hoddinott, J. (1994). A model of migration and remittances applied to western kenya. *Oxford Economic Papers*, 459–476.

Ilahi, N. and S. Jafarey (1999). Guestworker migration, remittances and the extended family: evidence from pakistan. *Journal of Development Economics 58*(2), 485–512.

LaLonde, R. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review*, 604–620.

Lucas, R. and O. Stark (1985). Motivations to remit: Evidence from botswana. *The Journal of Political Economy*, 901–918.

McKenzie, D., J. Gibson, and S. Stillman (2006). How important is selection? experimental vs non-experimental measures of the income gains from migration.

Osili, U. (2004). Migrants and Housing Investments: Theory and Evidence from Nigeria. *Economic Development and Cultural Change 52*(4), 821–849.

Osili, U. (2007). Remittances and savings from international migration: Theory and evidence using a matched sample. *Journal of Development Economics 83*(2), 446–465.

Rapoport, H. and F. Docquier (2006). The economics of migrants' remittances. *Handbook on the Economics of Giving, Reciprocity and Altruism 2*, 1135–1198.

Roger, G. (2010). Moral hazard with soft information.

Roger, G. (2012). Optimal Contract Under Moral Hazard with Soft Information. *UNSW Australian School of Business Research Paper* (2012ECON12).

Seror, M. (2012). Measuring Information Asymmetries and Modeling their Impact on Senegalese Migrants' Remittances. Master's thesis, Paris School of Economics and UC Berkeley.

Stark, O. (1999). *Altruism and beyond: An economic analysis of transfers and exchanges within families and groups*. Cambridge Univ Pr.

World Bank (2011, July). News & Broadcast: Migrants and Remittances.

Yang, D. (2011). Migrant remittances. *The Journal of Economic Perspectives 25*(3), 129–151.

Appendices

Table 24: Modalities of categorical housing variables

Name	Modalities
House type (quality)	1. Hut
	2. Shanty
	3. House, one block
	4. House, several blocks
	5. Multiple–storied house
	6. Building
Roof material (quality)	1. Thatch
	2. Tin
	3. Tile or slate
	4. Concrete

Table 25: Modalities of ordinal contact variables

Name	Modalities
Frequency of m's visits to h	0. Never went back
	1. Every 5 years
	2. Every 2–3 years
	3. Once a year
	4. Several times a year
	5. Many times a year (Mauritanian sample only)
	6. Twice a year (")
	7. Every 4–5 months (")
	8. Every 2–3 months (")
	9. Once a month (")
Frequency of m and h's phone contacts	0. No contact
	1. Once a year
	2. Twice a year
	3. Four times a year
	4. Once a month
	5. Once a week
Frequency of contacts with migrants from the same primary recipient household	0. Rarely
	1. Once a year
	2. Twice a year
	3. Once a month
	4. Once a week

Table 26: Correlation between t and X_{hm} (Mauritanian sample)

	(1) Total transfers (€)
fridge	54.05
	(77.85)
freezer	150.2
	(122.9)
dvd	118.8*
	(67.19)
motorbike	-16.27
	(80.87)
bike	52.64
	(74.95)
cart	87.26
	(79.59)
House type (quality)	9.267
	(67.79)
Roof material (quality)	47.53
	(44.42)
Nb of rooms	31.41**
	(14.07)
h owns their dwelling	-10.17
	(139.4)

Standard errors in parentheses

Bivariate regressions of total transfers in the past 12 months on *m*'s report of *X*.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 27: Correlation between t and X_{hm} (European sample)

	(1) Total transfers (€)
fridge	-37.30
	(123.5)
freezer	199.5
	(173.4)
tvset	201.8**
	(90.91)
dvd	228.1**
	(111.3)
radio	176.5***
	(64.14)
cd	109.7
	(77.42)
hoe	225.7*
	(123.9)
car	-75.58
	(102.9)
mopedscooter	-104.6
	(127.8)
motorbike	227.1
	(239.0)
bike	114.0
	(88.63)
canoe	-70.62
	(188.7)
House type (quality)	66.94
	(73.26)
Roof material (quality)	-4.510
	(117.6)
Nb of rooms	29.50
	(26.92)
h owns their dwelling	960.9***
	(224.6)

Standard errors in parentheses

Bivariate regressions of total transfers in the past 12 months on *m*'s report of *X*.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$