

# The Reliability of Voluntary Disclosures: Evidence from Hedge Funds

## Internet Appendix

Appendix A The Consolidated Hedge Fund Database .....	2
Appendix B Strategy Mappings .....	3
Table A.1 Listing of Vintage Dates .....	4
Table A.2 Summary Statistics for Lifetime Variables .....	5
Table A.3 Probit Regression for Additions .....	6
Table A.4 Probit Regression for Deletions .....	7
Table A.5 Probit Regression on Any Changes – Robustness Checks .....	8
Table A.6 Characteristics of the Reviser and Non-reviser funds.....	10
Figure A.1 Differences between “True” and Initial Returns .....	11
Figure A.2 Predicted Probabilities for Multinomial Logit on Revision Direction.....	12
Figure A.3 Portfolio Performance – Revisers and Non-revisers Single Database Check .....	13
Figure A.4 Portfolio Performance – Revisers and Non-revisers Median Return Check .....	14
Figure A.5 Portfolio Performance – Revisers and Non-revisers Recency Robustness Check .....	15

Update: 04 November 2011

Filename: AppendixHedgeFundRevisions04Nov2011.docx

## **Appendix A**

### **The Consolidated Hedge Fund Database**

As hedge funds can report to one or more databases, the use of any single source will fail to capture the complete universe of hedge fund data. We therefore aggregate data from TASS, HFR, CISDM, BarclayHedge and Morningstar, which together have 74,742 records of fund entries that comprise administrative information as well as returns and AUM data for hedge funds, fund of funds and CTAs. However this number hides the fact that there is significant duplication of information, as multiple providers often cover the same fund. To identify all unique entities, we must therefore consolidate the aggregated data. To do so, we adopt the following steps:

1. **Group the Data:** Records are grouped based on reported management company names. To do so, we first create a 'Fund name key' and a 'Management company key' for each data record, by parsing the original fund name and management company name for punctuations, filler words (e.g., 'Fund', 'Class'), and spelling errors. We then combine the fund and management name keys into 8,390 management company groups.
2. **De-Duplication:** Within a management company group, records are compared based on returns data (converted into US dollars), and 27,395 match sets are created out of matching records, allowing for a small error tolerance limit (10% deviation) to allow for data reporting errors.
3. **Selection:** Once all matches within all management company groups are identified, a single record representing the unique underlying fund is created for each match set. We pick the record with the longest returns data history available is selected from the match set, and fill in any missing administrative information using the remaining records in the match set. The process thus yields 27,395 representative funds.

We filter the fund data in a few ways to ensure data integrity. For example, removing return outliers and quarterly reporting funds, and ensuring funds have sufficient return or asset information. We also remove the Morningstar funds, given less than a third passed these quality filters, to ensure sufficient depth by database. The result is 18,382 funds.

## Appendix B Strategy Mappings

This table shows the broad strategies to which the underlying source strategies of the database vendors, HFR, TASS, CISDM and BarclayHedge, are mapped to. Examples of strategies are shown in the second column. The full set of more than 600 mappings is not shown. We also make use of fund type in the source database to aid in allocating an appropriate mapping. For example, a CTA with a source strategy dubbed Other will be allocated to the Managed Futures strategy with the other CTAs and not into the Other hedge fund category.

---

<b>Mapped Strategy</b>	<b>Examples of source strategies</b>
Security Selection	Equity Long/Short, Equity Arbitrage, Equity Long/Short - Growth Bias, Equity Market Neutral, Equity Market Neutral - US Value Long/Short
Macro	Global Macro, Global Macro - FX only, Global Macro - Quantitative, Macro - Active Trading
Relative Value	Merger Arbitrage, Equity Market Neutral - Relative Value, Single Strategy - Event Driven Risk Arbitrage, Statistical Arbitrage
Directional Traders	Dedicated Short Bias, Equity Long Only, Equity Long/Short - Long biased, Market Timing, Single Strategy - Tactical trading
Fund-of-Funds	(By fund type) , Fund of Funds, Fund of Funds - Strategic, Conservative - Absolute Return Fund of Funds, Fund of Funds - Nondirectional, Fund of Funds - Derivatives
Multi-Process	Multi-process, Multi Strategy - Arbitrage, Equity Hedge - Multi-Strategy, Event Driven Multi Strategy
Emerging	Emerging Markets, Emerging Markets - Central Asia focus, Equity Long/Short - Emerging Markets, Emerging Markets - Directional, Emerging Markets - Global
Fixed Income	Convertible Arbitrage, Fixed Income - Arbitrage, Fixed Income - ABS/Sec. Loans, Fixed Income - Structured Credit, Global Debt, Distressed Securities - Stressed High Yield Bonds
Other	Other, Undefined, Closed-end funds
Managed Futures	(By CTA fund type), Managed Futures, Global trend, Discretionary - CTA Managed Futures, Systematic - Systematic arbitrage & counter-trend

---

**Table A.1**  
**Listing of Vintage Dates**

This table shows the vintage dates of the 40 snapshots.

---

<b>Number</b>	<b>Vintage date</b>
1	Jul 2007
2	Jan 2008
3	Feb 2008
4	Mar 2008
5	Apr 2008
6	May 2008
7	Jun 2008
8	Jul 2008
9	Aug 2008
10	Sep 2008
11	Oct 2008
12	Nov 2008
13	Dec 2008
14	Jan 2009
15	Mar 2009
16	Apr 2009
17	May 2009
18	Jun 2009
19	Jul 2009
20	Aug 2009
21	Sep 2009
22	Oct 2009
23	Dec 2009
24	Jan 2010
25	Feb 2010
26	Mar 2010
27	Apr 2010
28	May 2010
29	Jun 2010
30	Jul 2010
31	Aug 2010
32	Sep 2010
33	Oct 2010
34	Nov 2010
35	Dec 2010
36	Jan 2011
37	Feb 2011
38	Mar 2011
39	Apr 2011
40	May 2011

---

**Table A.2**  
**Summary Statistics for Lifetime Variables**

This table shows for the sample of funds the Lifetime Assets and Return Averages, Std Deviations and Medians. LIFEN is the number of returns the fund reported. RHO1 is return first autocorrelation. (Figures are unwinsorised in this table and taken from the last vintage.)

	AUM Avg	AUM Std	AUM Median	Return Avg	Return Std	Return Median	RHO1	LIFEN
Observations	18,382	18,382	18,382	18,382	18,382	18,382	18,382	18,382
Mean	149,289,134	79,707,015	135,439,957	0.644	4.102	0.677	0.139	66.422
Std dev	1,491,667,969	744,463,251	1,413,991,918	1.180	3.638	1.008	0.222	45.342
99th perc	1,723,491,752	972,471,117	1,595,549,937	4.652	18.278	3.770	0.655	207
75th perc	73,538,781	35,962,608	64,020,000	1.008	5.152	1.020	0.284	88
Median	22,754,853	9,070,742	19,444,848	0.552	3.032	0.610	0.139	54
25th perc	5,891,644	2,018,060	4,574,000	0.181	1.813	0.240	0.005	32
1 perc	101,520	-	-	2.283	0.437	2.047	0.415	13

**Table A.3**  
**Probit Regression for Additions**

The table shows the marginal effects from a probit regression. The dependent variable is the dummy reflecting whether a fund had an Addition over the period of all the vintages. This is explained by the rank of lifetime variables of average assets under management, average return, return standard deviation, return first auto correlation ( $\rho_1$ ) and the number of returns the fund reported (lifen). Other relevant fund variables are an offshore dummy, total restrictions variable (measured as the sum of the reported lockup periods) and an audit information flag. Relevant control dummies of fund strategy and database of fund are included. Regressors are described in the text.  $dF/dx$  is for discrete change of dummy variable from 0 to 1, and the slope at the mean for continuous variables. Standard errors estimated by clustering by database. The number of stars \* denote significance at 10%, 5% and 1% respectively.

<b>Additions</b>	<b>dF/dx</b>	<b>Mean</b>	<b>Robust SE</b>	<b>z</b>	
lifeaumavgrank	-0.002	0.500	0.001	-1.760	*
liferetavgrank	-0.004	0.500	0.006	-0.670	
liferetstdrank	0.006	0.500	0.004	1.450	
rho1rank	0.003	0.500	0.004	0.740	
lifen	0.000	66.422	0.000	6.020	***
offshore	0.001	0.501	0.002	0.380	
lockup	0.000	164.623	0.000	0.580	
audit	0.010	0.712	0.004	1.980	**
<i>DB HFR</i>	-0.006	0.258	0.001	-4.700	***
<i>DB CISDM</i>	-0.013	0.092	0.001	-5.430	***
<i>DB BarclayHedge</i>	-0.003	0.290	0.001	-3.710	***
<i>Macro</i>	-0.004	0.065	0.003	-1.060	
<i>Relative Value</i>	0.003	0.014	0.008	0.430	
<i>Directional Traders</i>	-0.004	0.128	0.004	-0.990	
<i>Fund of Funds</i>	0.007	0.264	0.002	3.840	***
<i>Multi-Process</i>	-0.004	0.102	0.001	-2.390	**
<i>Emerging</i>	0.002	0.045	0.002	1.390	
<i>Fixed Income</i>	0.005	0.052	0.009	0.650	
<i>Other</i>	0.043	0.009	0.007	11.040	***
<i>Managed Futures</i>	0.004	0.157	0.004	1.030	
Number observations			18,382		
Log pseudolikelihood			-1,647.63		
Pseudo R2			9.04%		

**Table A.4**  
**Probit Regression for Deletions**

The table shows the marginal effects from a probit regression. The dependent variable is the dummy reflecting whether a fund had a Deletion over the period of all the vintages. This is explained by the rank of lifetime variables of average assets under management, average return, return standard deviation, return first auto correlation ( $\rho_1$ ) and the number of returns the fund reported (lifen). Other relevant fund variables are an offshore dummy, total restrictions variable (measured as the sum of the reported lockup periods) and an audit information flag. Relevant control dummies of fund strategy and database of fund are included. Regressors are described in the text.  $dF/dx$  is for discrete change of dummy variable from 0 to 1, and the slope at the mean for continuous variables. Standard errors estimated by clustering by database. The number of stars \* denote significance at 10%, 5% and 1% respectively.

<b>Deletions</b>	<b>dF/dx</b>	<b>Mean</b>	<b>Robust SE</b>	<b>z</b>	
lifeaumavgrank	0.013	0.500	0.005	2.430	**
liferetavgrank	-0.030	0.500	0.025	-1.170	
liferetstdrank	0.009	0.500	0.005	1.730	*
rho1rank	-0.005	0.500	0.012	-0.460	
lifen	0.000	66.422	0.000	19.050	***
offshore	0.019	0.501	0.007	2.850	***
lockup	0.000	164.623	0.000	-0.620	
audit	0.018	0.712	0.003	6.170	***
<i>DB HFR</i>	-0.007	0.258	0.002	-3.880	***
<i>DB CISDM</i>	-0.031	0.092	0.002	-16.320	***
<i>DB BarclayHedge</i>	-0.021	0.290	0.002	-10.230	***
<i>Macro</i>	0.004	0.065	0.006	0.810	
<i>Relative Value</i>	0.050	0.014	0.015	4.070	***
<i>Directional Traders</i>	0.006	0.128	0.004	1.390	
<i>Fund-of-Funds</i>	0.022	0.264	0.003	7.090	***
<i>Multi-Process</i>	-0.011	0.102	0.004	-2.400	**
<i>Emerging</i>	0.019	0.045	0.008	2.650	***
<i>Fixed Income</i>	0.015	0.052	0.015	1.080	
<i>Other</i>	0.017	0.009	0.021	0.900	
<i>Managed Futures</i>	0.008	0.157	0.006	1.410	
Number observations			18,382		
Log pseudolikelihood			-3,931.17		
Pseudo R2			4.19%		

**Table A.5**  
**Probit Regression on Any Changes – Robustness Checks**

As per Table III, these are results of the probit regressions on any changes, but are showing the marginal changes estimates at different quantile ranks, rather than the mean for the continuous ranked variables.

Panel A. Marginal effects of ranks at 0.75

<b>Change</b>	<b>dF/dx</b>	<b>Mean</b>	<b>Robust SE</b>	<b>z</b>	
lifeaumavgrank	0.245	0.750	0.053	4.590	***
liferetavgrank	-0.094	0.750	0.056	-1.670	*
liferetstdrank	0.065	0.750	0.043	1.510	
rho1rank	0.121	0.750	0.015	8.210	***
lifef	0.002	66.297	0.000	4.870	***
offshore	-0.009	0.501	0.007	-1.240	
lockup	0.000	164.591	0.000	5.170	***
audit	0.183	0.713	0.097	1.890	*
<i>DB HFR</i>	-0.017	0.259	0.009	-1.810	*
<i>DB CISDM</i>	-0.069	0.092	0.078	-0.890	
<i>DB BarclayHedge</i>	0.106	0.290	0.011	9.620	***
<i>Macro</i>	0.086	0.065	0.007	12.110	***
<i>Relative Value</i>	0.183	0.014	0.055	3.330	***
<i>Directional Traders</i>	-0.006	0.128	0.014	-0.420	
<i>Fund-of-Funds</i>	0.219	0.263	0.016	13.970	***
<i>Multi-Process</i>	0.059	0.102	0.017	3.540	***
<i>Emerging</i>	0.121	0.045	0.011	11.100	***
<i>Fixed Income</i>	0.026	0.053	0.031	0.850	
<i>Other</i>	0.123	0.010	0.111	1.110	
<i>Managed Futures</i>	0.120	0.157	0.042	2.850	***



Panel B. Marginal effects of ranks at 0.25

<b>Change</b>	<b>dF/dx</b>	<b>Mean</b>	<b>Robust SE</b>	<b>z</b>	
lifeaumavgrank	0.218	0.250	0.045	4.870	***
liferetavgrank	-0.084	0.250	0.053	-1.590	
liferetstdrank	0.058	0.250	0.036	1.600	
rho1rank	0.108	0.250	0.013	8.320	***
lifen	0.002	66.297	0.000	4.170	***
offshore	-0.008	0.501	0.006	-1.270	
lockup	0.000	164.591	0.000	5.690	***
audit	0.156	0.713	0.079	1.980	**
<i>DB HFR</i>	-0.015	0.259	0.008	-1.960	**
<i>DB CISDM</i>	-0.060	0.092	0.065	-0.910	
<i>DB BarclayHedge</i>	0.097	0.290	0.012	7.760	***
<i>Macro</i>	0.080	0.065	0.007	12.180	***
<i>Relative Value</i>	0.179	0.014	0.060	2.990	***
<i>Directional Traders</i>	-0.005	0.128	0.012	-0.420	
<i>Fund-of-Funds</i>	0.207	0.263	0.013	16.410	***
<i>Multi-Process</i>	0.054	0.102	0.017	3.200	***
<i>Emerging</i>	0.114	0.045	0.012	9.190	***
<i>Fixed Income</i>	0.024	0.053	0.028	0.850	
<i>Other</i>	0.117	0.010	0.114	1.030	
<i>Managed Futures</i>	0.111	0.157	0.041	2.710	***

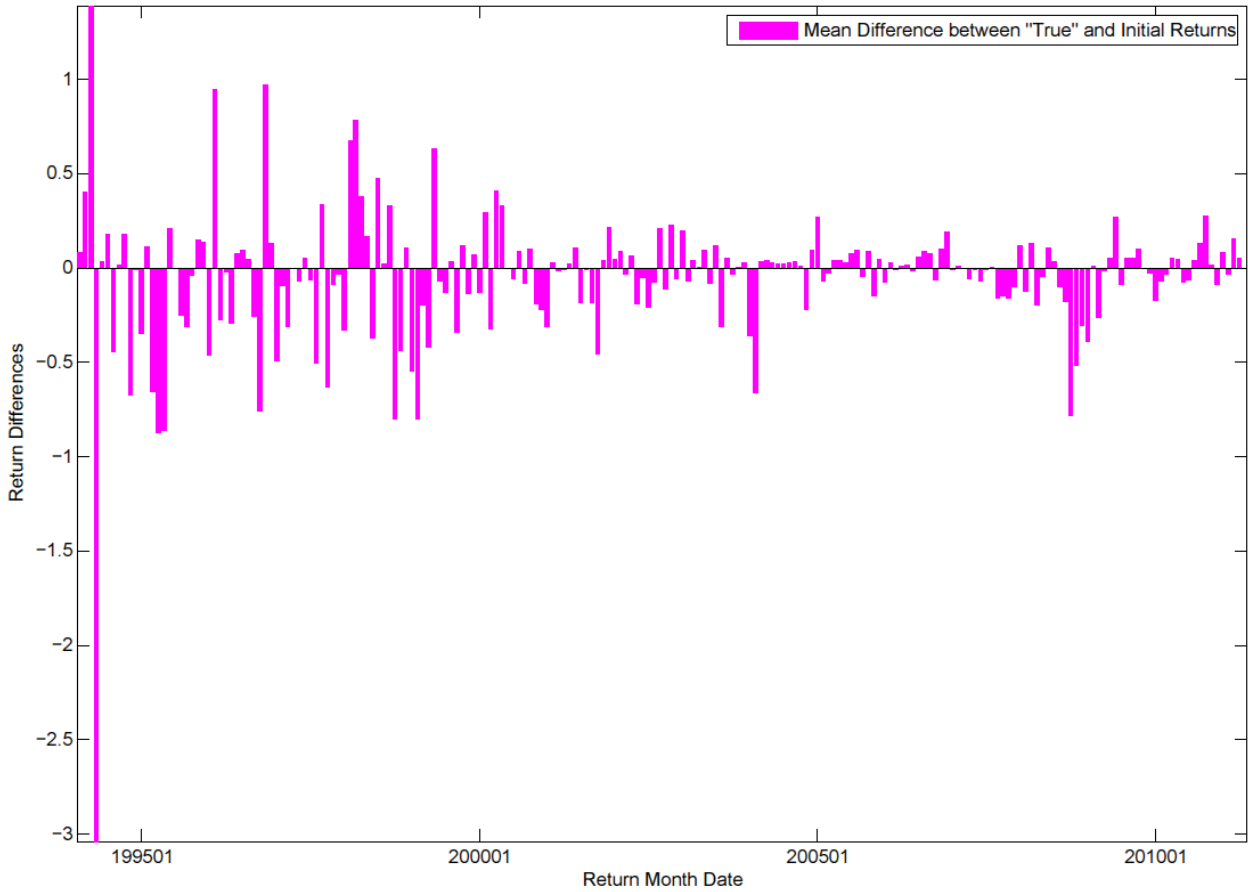
**Table A.6**  
**Characteristics of the Reviser and Non-reviser funds**

This table shows the differences in characteristics between the reviser and non-reviser groups of funds using the status of the funds at the last vintage. The non-reviser funds at this stage have never revised between vintages. Once a fund revises a return it joins the reviser portfolio and it stays out of the non-reviser group. Lifetime AUM and return measures are used for the funds, not the period in which they belonged to the group. There are 11,476 non-reviser funds out of the 18,382 funds. *t*-statistics of the differences between groups assume a common variance.

<b>Variable</b>	<b>Revisers</b>		<b>Non-revisers</b>		<b><i>t</i>-stat diff</b>	<b><i>p</i>-value</b>
	<b>Mean</b>	<b>Std Dev</b>	<b>Mean</b>	<b>Std Dev</b>		
Lifetime AUM Average \$m	180.91	1,479.51	130.26	1,498.68	2.230	0.026
Lifetime Return Average	0.636	0.987	0.649	1.282	-0.680	0.497
Rho1	0.186	0.218	0.111	0.219	22.508	0.000
Return count	79.62	50.41	58.48	39.95	31.420	0.000
Total lock	198.416	261.719	144.288	213.943	15.251	0.000

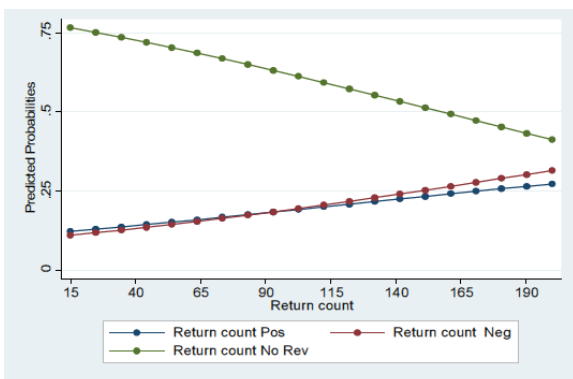
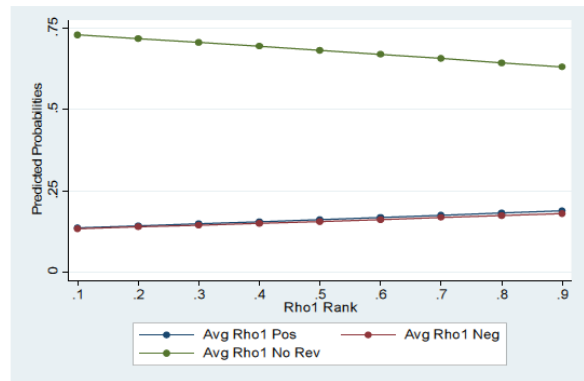
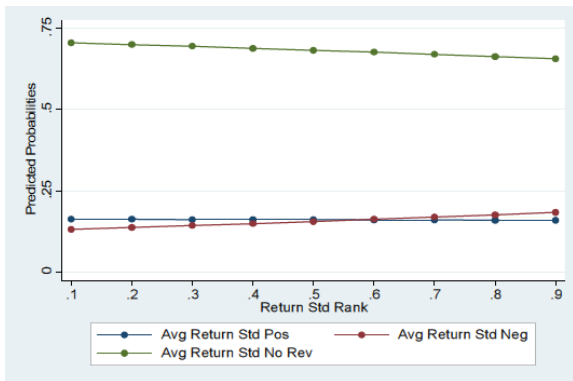
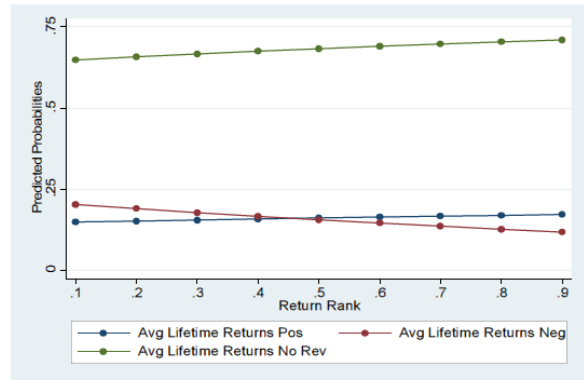
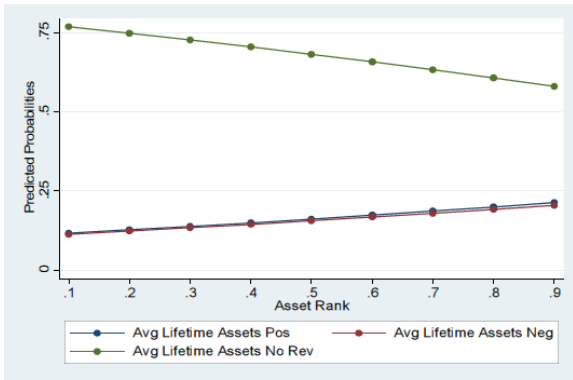
**Figure A.1**  
**Differences between “True” and Initial Returns**

This figure shows the average return differences between the last expression of the return at the most recent available database (denoted “True”) and the first time the return is expressed in a database (denoted Initial). Significant differences only are shown (so zero differences and minor differences due to changes in expression of significant digits for the same return value are excluded). [This is averaging over all differences unlike the separation by sign in Figure IV]



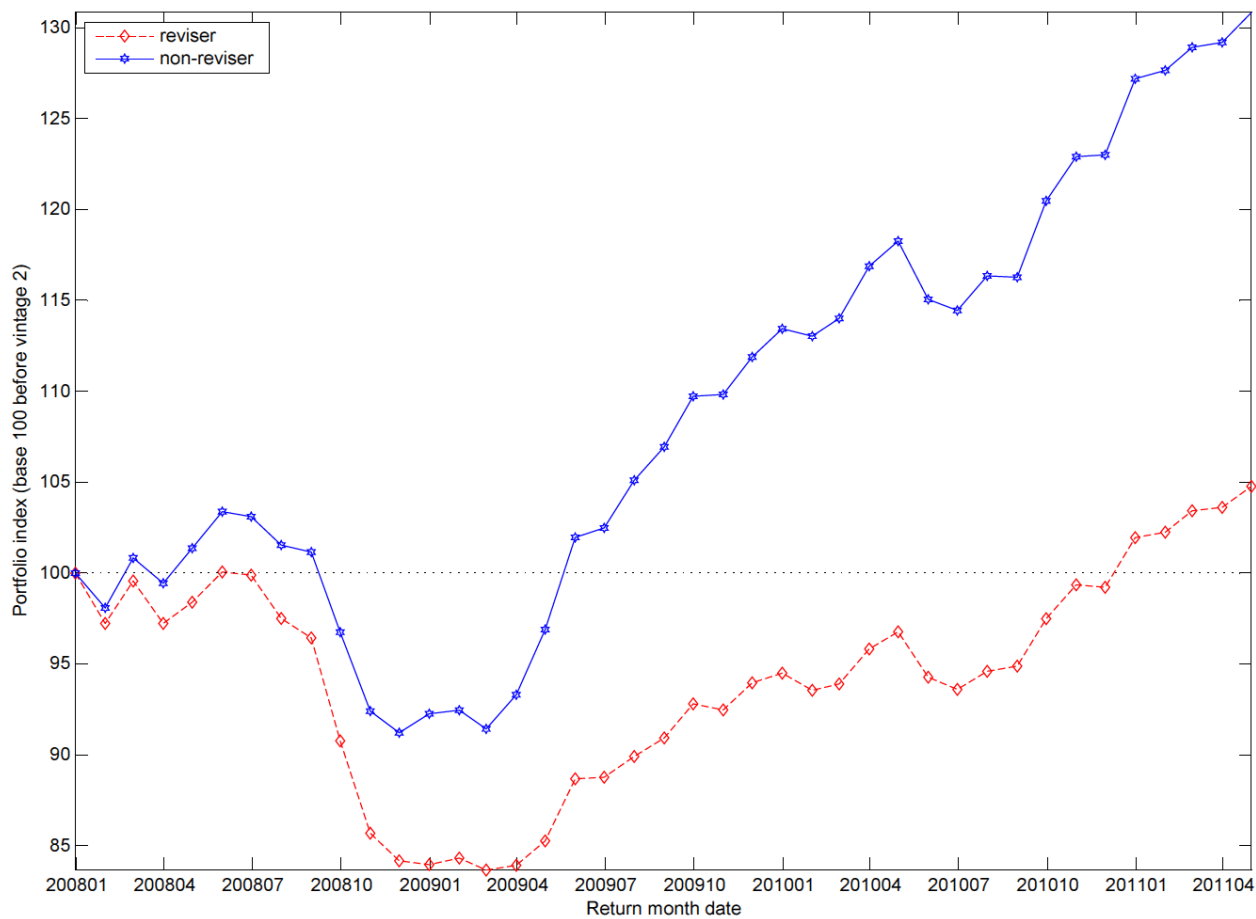
## Figures A.2 Predicted Probabilities for Multinomial Logit on Revision Direction

These figures show the predicted probabilities for the multinomial logit regression in Table VII. Variables are kept at their mean values except for the variable depicted in the x axis which varies from 10<sup>th</sup> to 90<sup>th</sup> percentile in value.



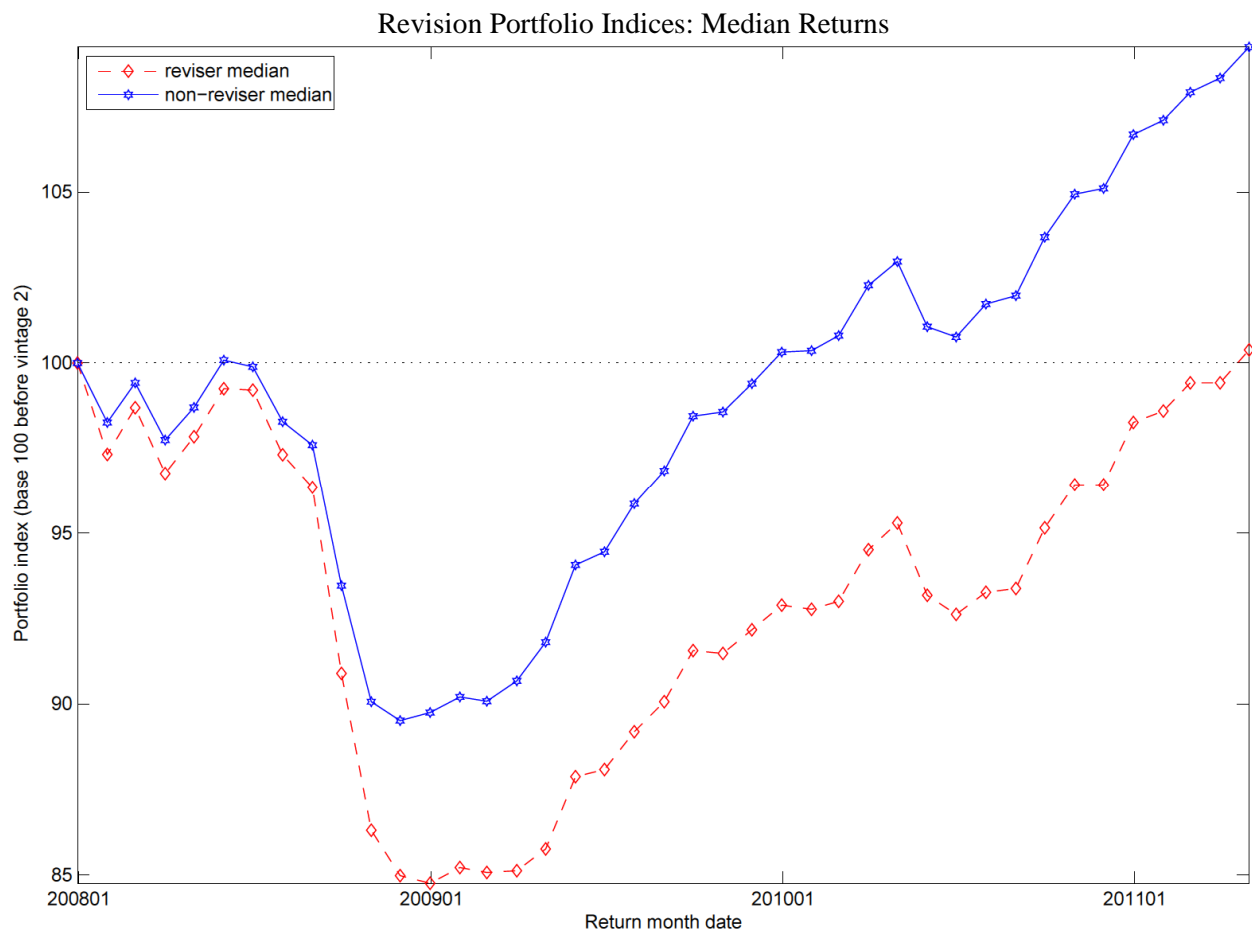
**Figure A.3**  
**Portfolio Performance – Revisers and Non-revisers Single Database Check**

The figure shows the cumulative performance of the reviser and non-reviser portfolios for a single database, in this case BarclayHedge. The non-reviser portfolio holds performance of funds that never revise between vintages plus the early records of funds before they become revisers. For example, if a fund first revises at vintage  $v$ ; its earlier performance will be included in the non-reviser portfolio as it had not yet been classified as a reviser. But once it joins the reviser portfolio it stays out of the non-reviser portfolio. The index is based to 100 at 31 December 2007, just before the second vintage starts. Returns are equally weighted in portfolios.



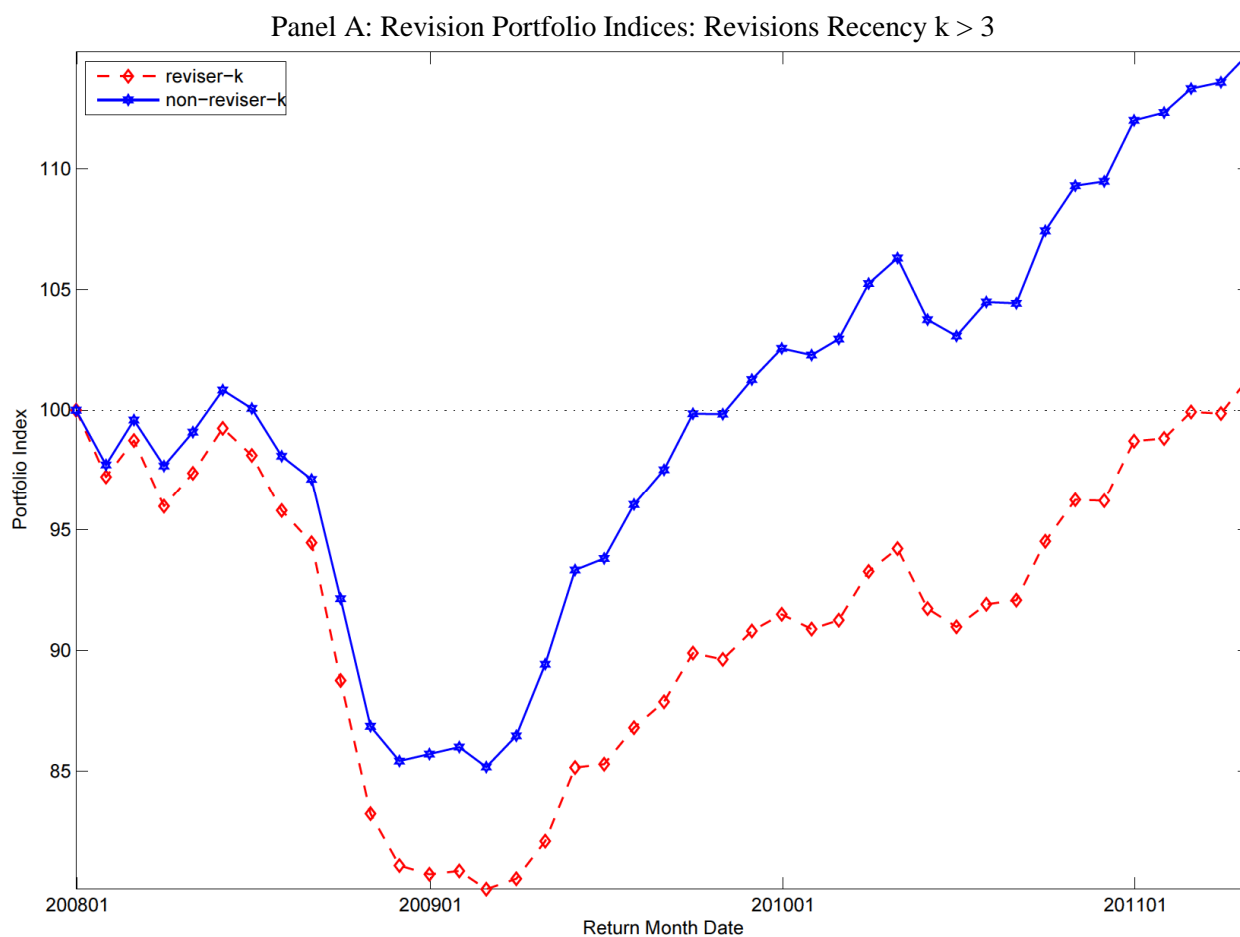
**Figure A.4**  
**Portfolio Performance – Revisers and Non-revisers Median Return Check**

The figure shows the cumulative performance of the reviser and non-reviser portfolios. The non-reviser portfolio holds performance of funds that never revise between vintages plus the early records of funds before they become revisers. For example, if a fund first revises at vintage  $v$ ; its earlier performance will be included in the non-reviser portfolio as it had not yet been classified as a reviser. But once it joins the reviser portfolio it stays out of the non-reviser portfolio. The index is based to 100 at 31 December 2007, just before the second vintage starts. Returns are the median returns of the portfolios.



**Figure A.5**  
**Portfolio Performance – Revisers and Non-revisers Recency Robustness Check**

The figure shows the cumulative performance of the reviser and non-reviser portfolios. The non-reviser portfolio holds performance of funds that never revise between vintages plus the early records of funds before they become revisers. For example, if a fund first revises at vintage  $v$ ; its earlier performance will be included in the non-reviser portfolio as it had not yet been classified as a reviser. But once it joins the reviser portfolio it stays out of the non-reviser portfolio. The index is based to 100 at 31 December 2007, just before the second vintage starts. Returns in this robustness check exclude revisions based on recency threshold  $k$  as explained in the paper. Panel A shows  $k > 3$  and Panel B  $k > 12$  months.



Panel B: Revision Portfolio Indices: Revisions Recency  $k > 12$

