Dynamics and Heterogeneity of Subjective Stock Market Expectations*

Florian Heiss (HHU Düsseldorf)
Michael Hurd (RAND, Santa Monica)
Maarten van Rooij (DNB†)
Tobias Rossmann (LMU Munich)
Joachim Winter (LMU Munich)

September 2019

*The full paper is available as Discusssion Paper No. 157 CRC/TR 190.

†The views expressed in this paper are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank (DNB).

Motivation

- Stock market returns were among the first applications of probabilistic expectation elicitation
- Several papers document
  - cross-sectional heterogeneity
  - predictive power for stock market participation
- Less is known about updating behavior (not only of stock market expectations)

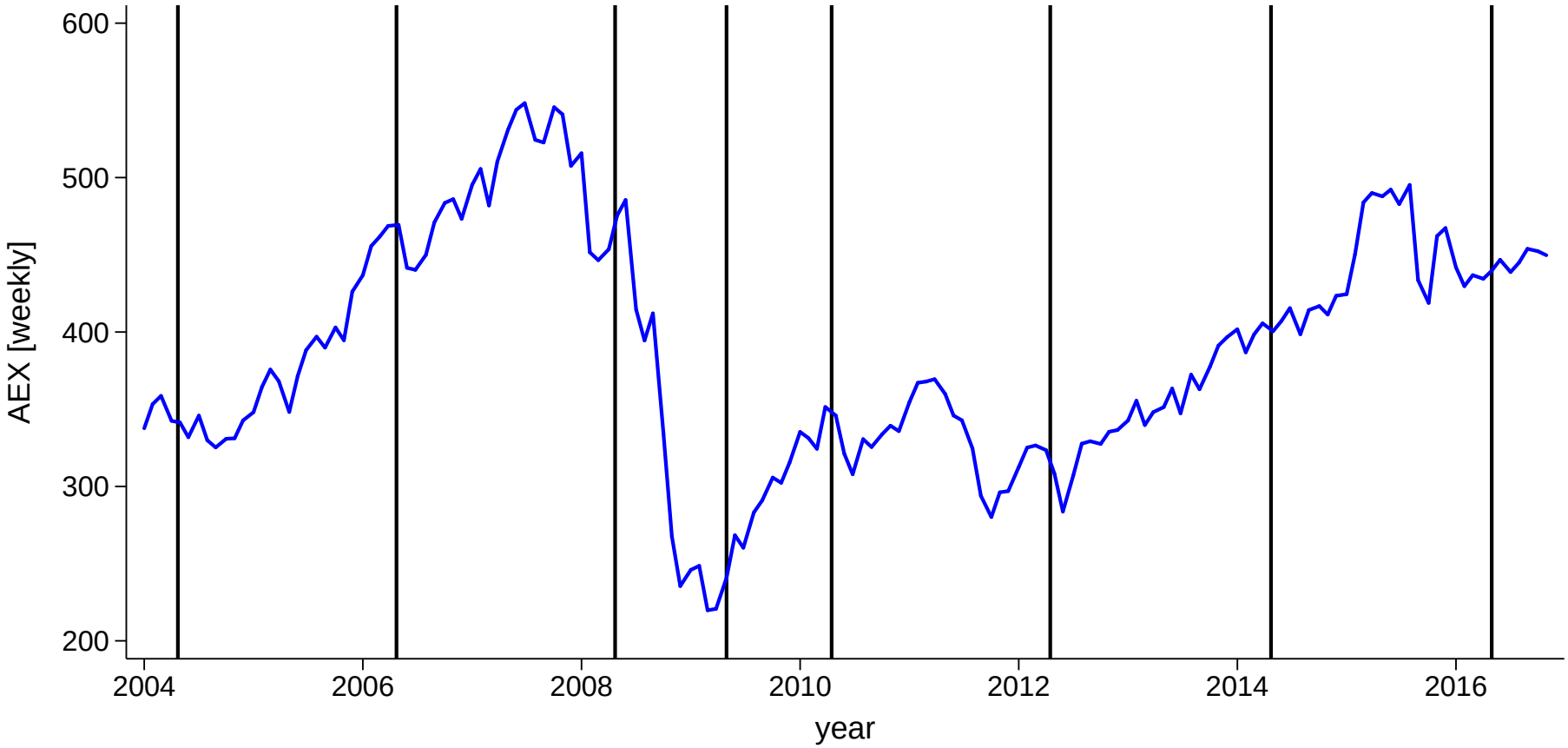
What we do

- Study updates of stock market expectations over a long period (2004–2016)
  - Data from a representative longitudinal survey (bi-annually)
  - Probabilistic elicitation with eight probabilities per measurement
  - Consistently used in all waves
- Estimate subjective mean and variance, for each person and measurement
- Develop and estimate a dynamic panel data model with a finite mixture of updating types, following Dominitz and Manski (2011, *JAE*)

Data I

- CentER Panel (Netherlands)
- Roughly 2,000 Dutch households
- Can be merged with DNB Household Survey (DHS) data
- Bi-annual module on stock market expectations, from 2004 to 2016
- Cross-sectional heterogeneity documented in Hurd, van Rooij, Winter (2011, *JAE*)

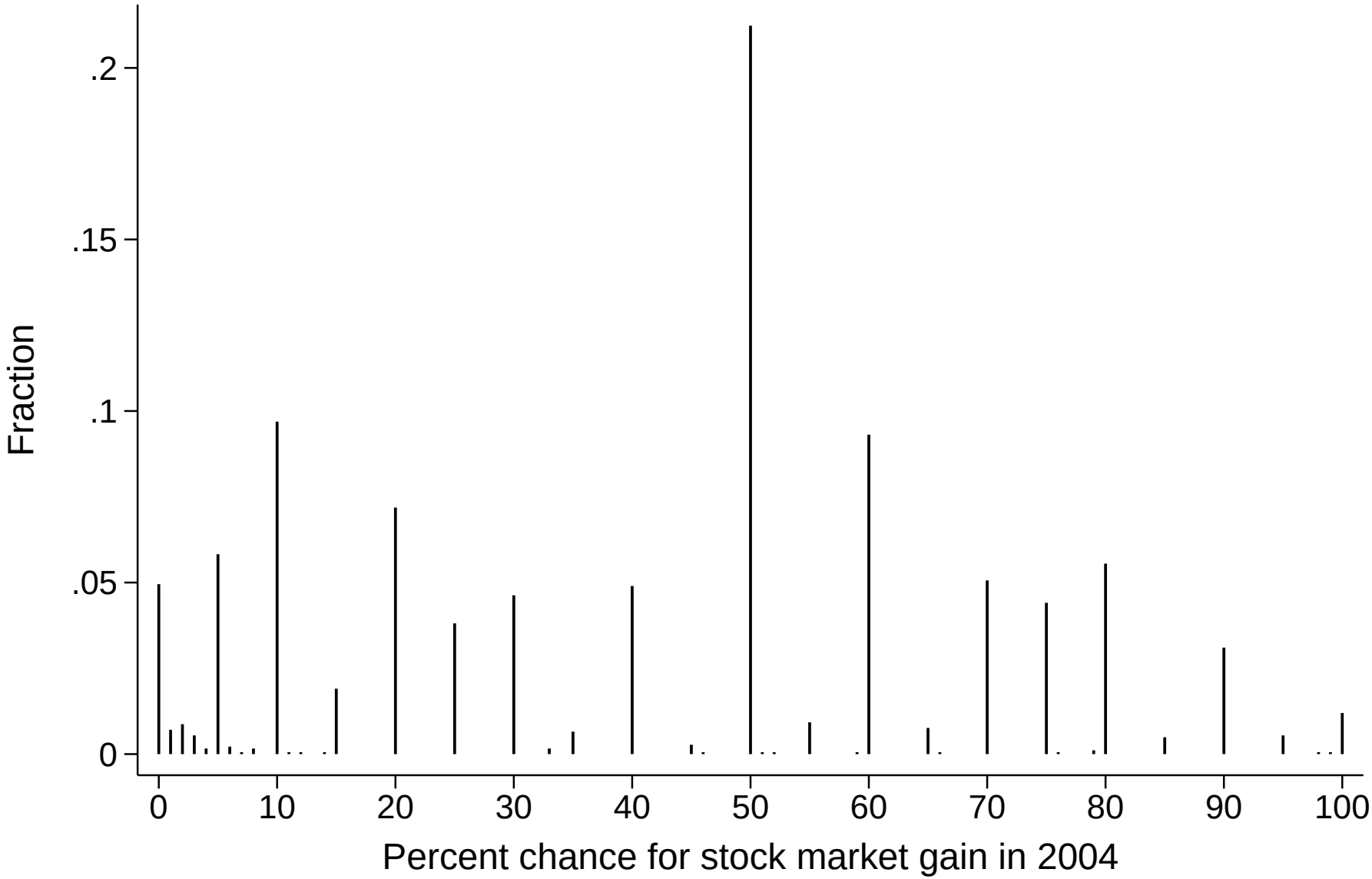
Data II


600
500
400
300
200
AEX [weekly]
2004
2006
2008
2010
2012
2014
2016
year

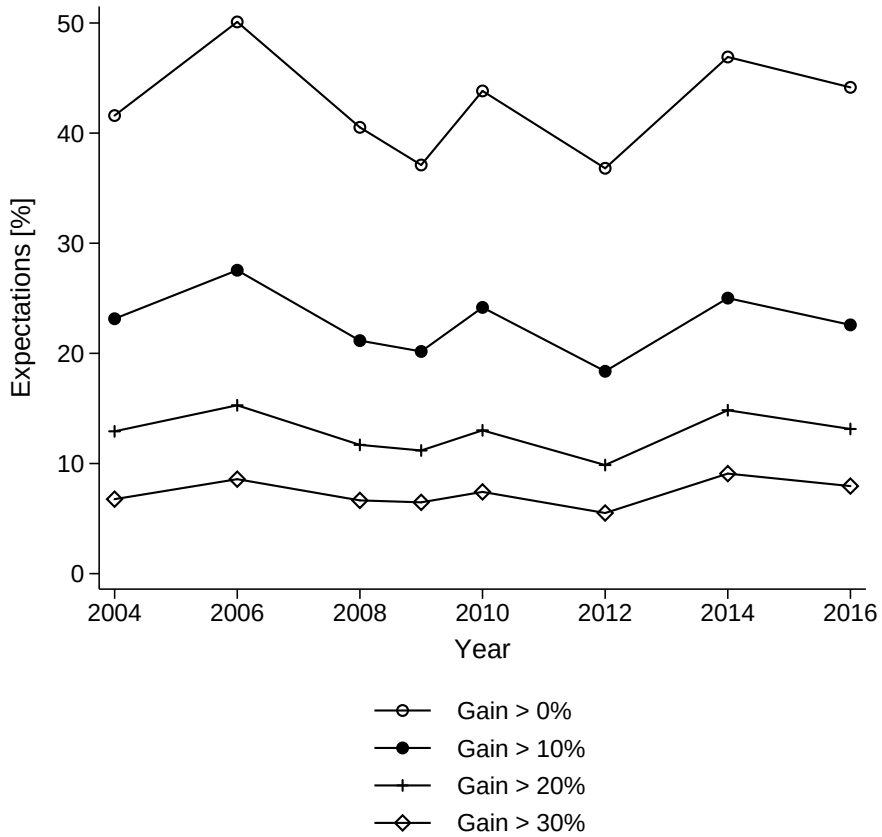
Stock market expectations

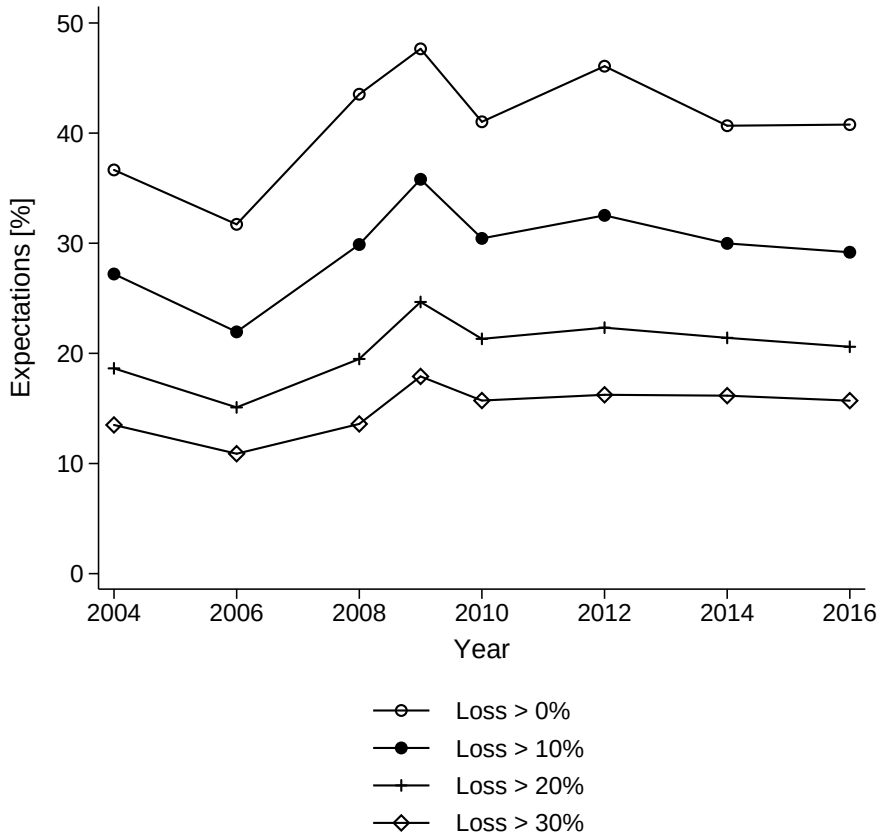
- By next year at this time, what is the percent chance that mutual fund shares invested in blue chip stocks like those in the Amsterdam Stock Exchange (AEX) will be worth more than they are today?
- Setting: Unexpectedly receive 10,000 EUR from rich relative
- In total eight questions:
  - Gain domain: 0%, 10%, 20% and 30%
  - Loss domain: -0%, -10%, -20% and -30%
- Randomization whether gain or loss domain is asked first

Example: Positive stock market return



Yearly averages of the eight responses





Updating types

Dominitz and Manski (2011, *JAE*)

- Population can be split into three different updating types:
  - **Random Walk (RW)**: returns are iid over time and historical records are used to predict future returns
  - **Persistence (P)**: recent market movements will persist in the near future
  - **Mean Reversion (MR)**: recent market movements will be reverted in the near future

Panel data model

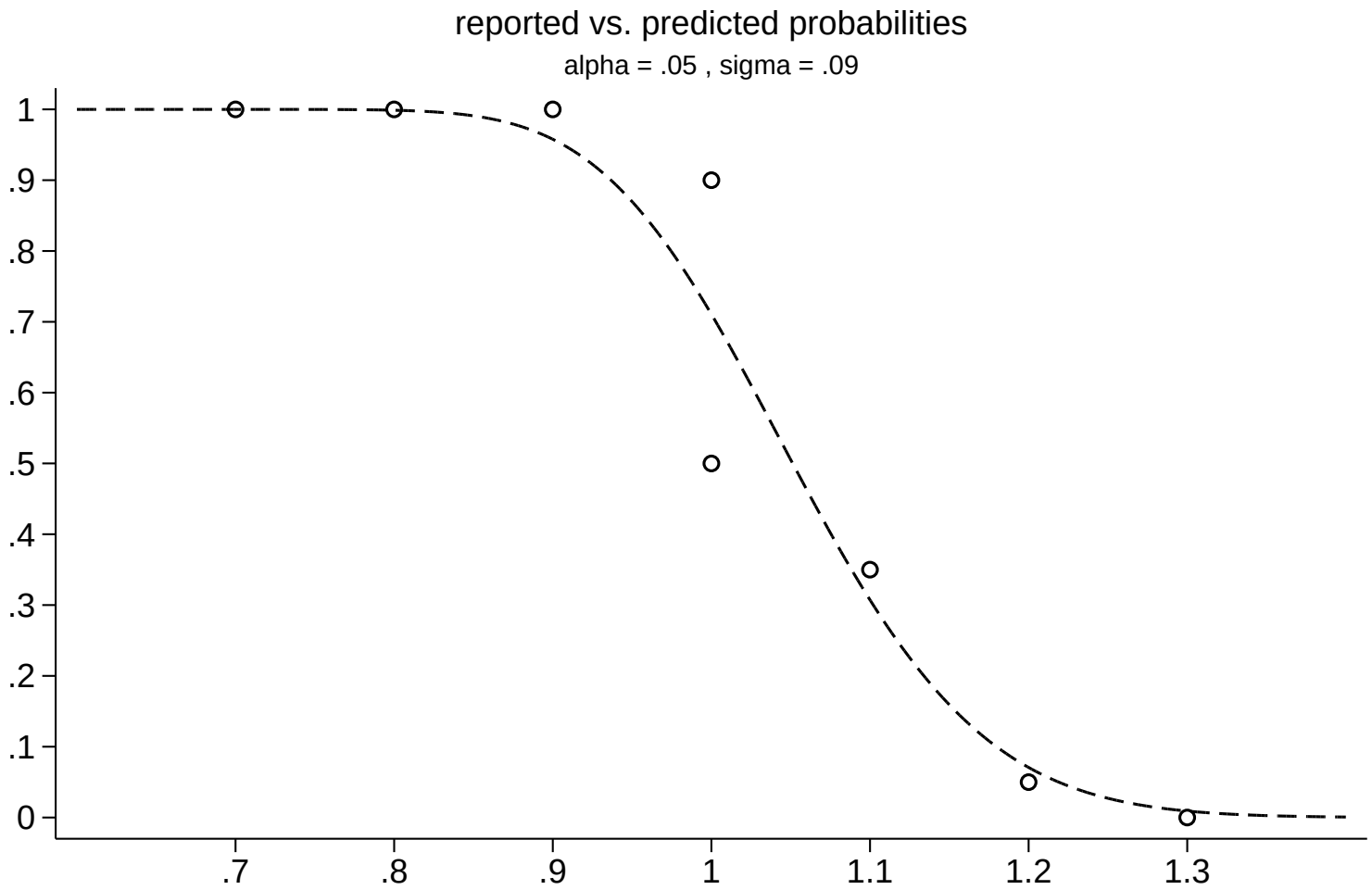
$$\mu_{it}^{*} = \gamma_k R_{t-1} + \boldsymbol{x}_{it}\boldsymbol{\beta} + \alpha_i + u_{it} \ , \qquad k \in \{1, 2, 3\} \tag{1}$$

- μ_{it}^{*}: mean of subjective stock market distribution
- R_{t-1}: AEX return from previous year
- Sign of γ corresponds to three latent classes:
  - $\gamma_1 = 0$: Random Walk (RW)
  - $\gamma_2 < 0$: Mean Reversion (MR)
  - $\gamma_3 > 0$: Persistence (P)
- $\boldsymbol{x}_{it}$: vector of potentially time-varying covariates
- α_i: individual-specific random effect
- u_{it}: idiosyncratic error
- Note: The $\gamma_1 = 0$ type also contains those that do not react to past returns for other reasons (observationally equivalent to "true" RW individuals)

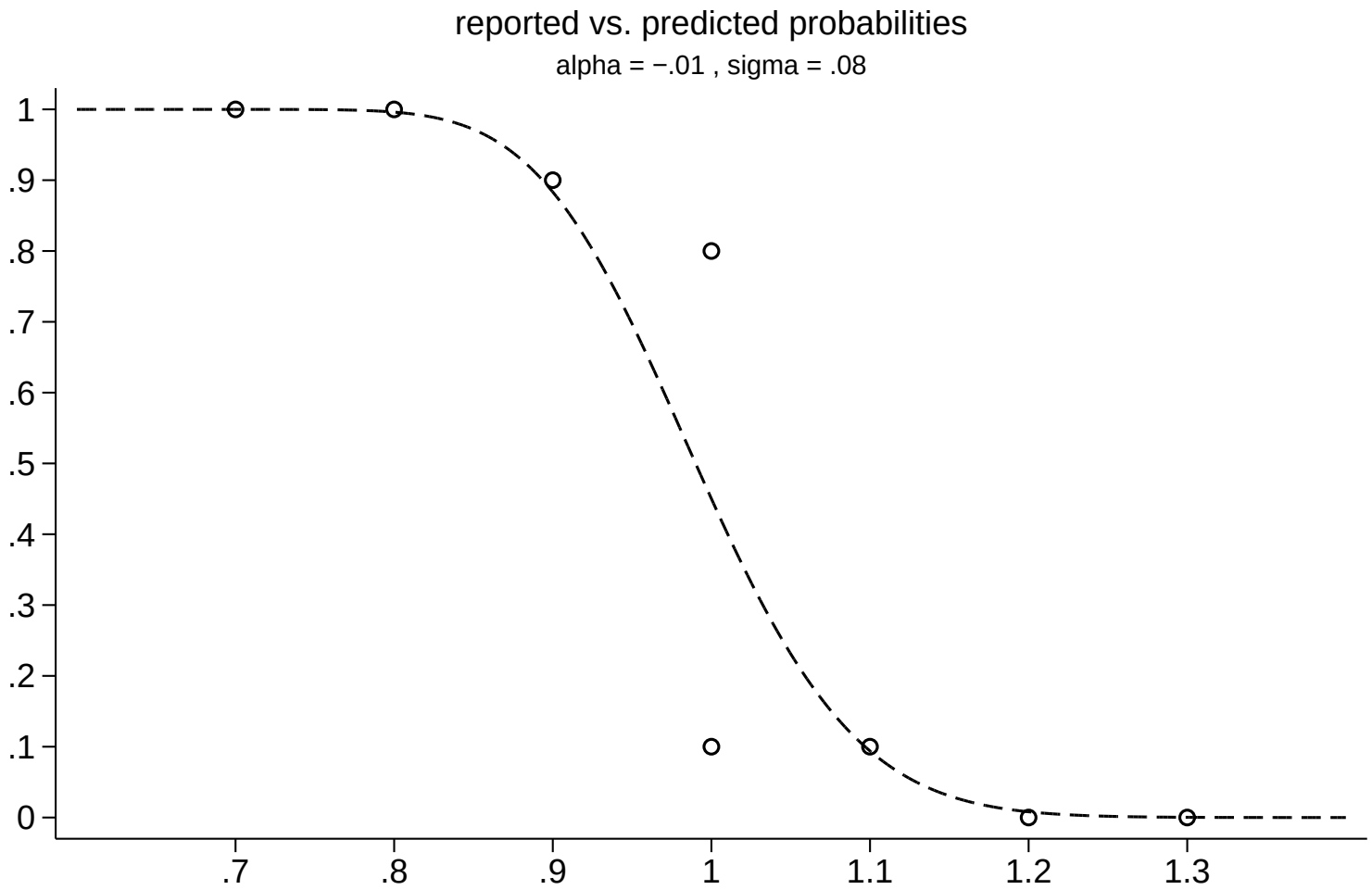
Additional features of the model I

- Type probabilities (for RW, MR and P):
  - Random effects multinomial logit (MNL) specification
  - May depend on covariates
- Construction of subjective means from the eight probability responses:
  - Fit a normal CDF through the eight points by Maximum Likelihood

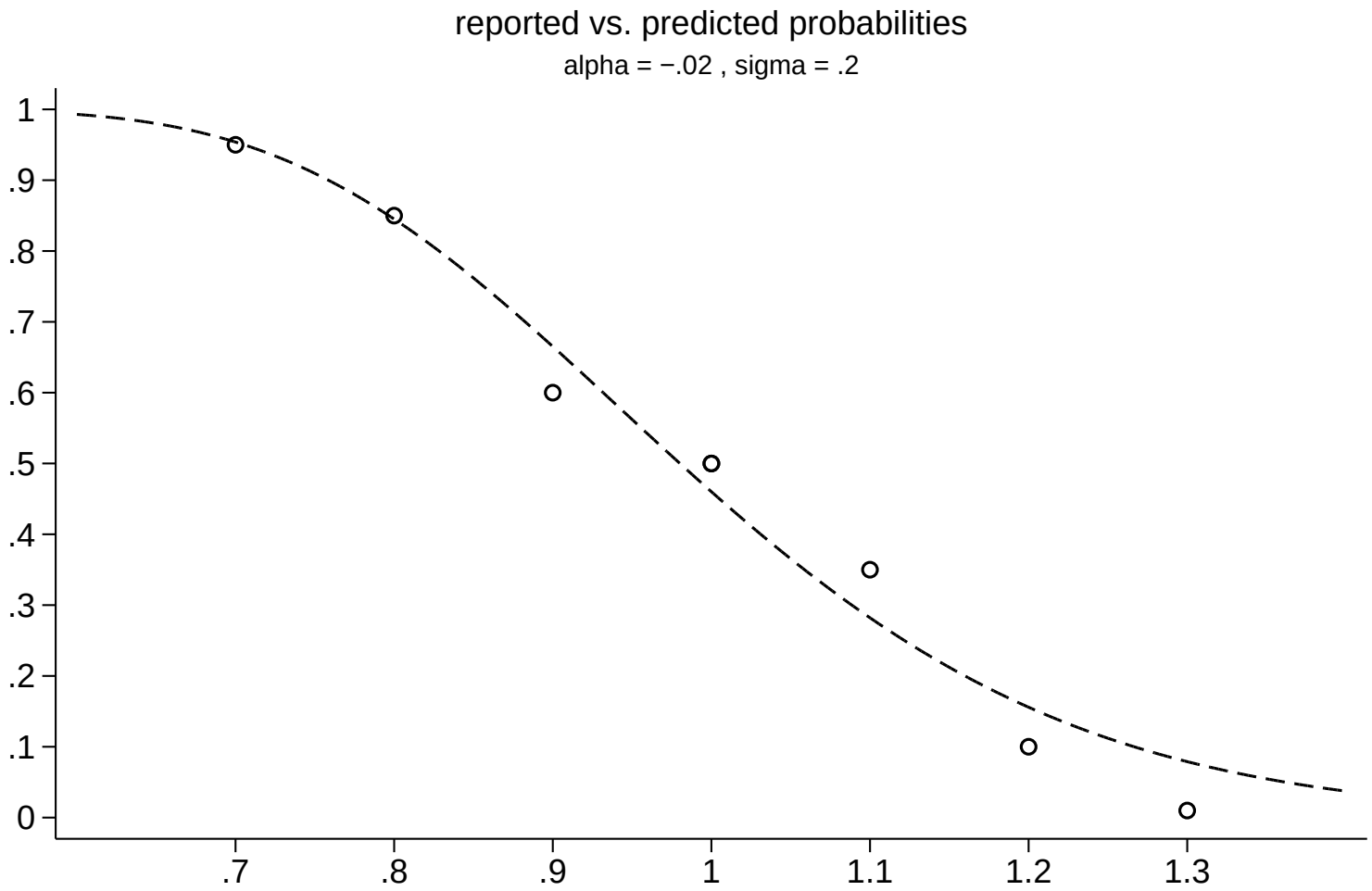
Examples of individual subjective distributions



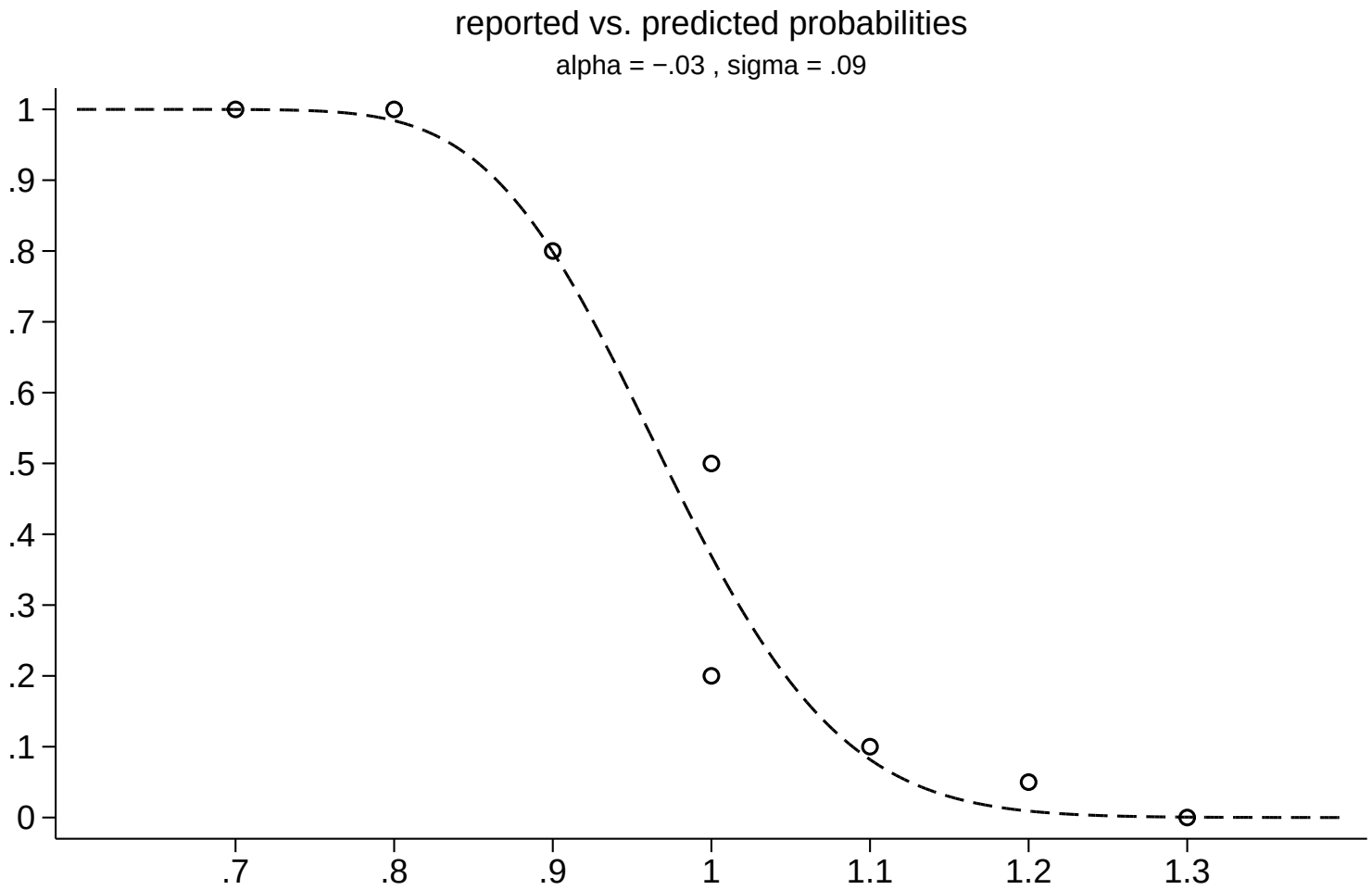
Examples of individual subjective distributions



Examples of individual subjective distributions



Examples of individual subjective distributions



Additional features of the model II

- Rounding
  - Following Kleinjans and van Soest (2014)
  - Three different rounding types (R1, R5 and R50)
  - Separate rounding for each of the eight questions
  - Random effects ordered probit specification
  - May depend on covariates

Estimation

- Joint estimation of the updating and rounding models
- Arbitrary var-cov structure for the individual-specific random effects
- Maximum Simulated Likelihood (MSL) to avoid numerical integration in several dimensions

Main equation: $\mu_{it}^* = \gamma_k R_{t-1} + \boldsymbol{x}_{it}\beta + \alpha_i + u_{it}$

	(1) Constants only		(2) Full model		(3) Restricted return coeff.	
γ_1: Return coeff. CI1	0.0271***	[0.0015]	0.0286***	[0.0014]		
γ_2: Return coeff. CI2	-0.5774***	[0.0167]	-0.5890***	[0.0166]	-0.5964***	[0.0185]
γ_3: Return coeff. CI3	0.6576***	[0.0125]	0.7143***	[0.0127]	0.6154***	[0.0104]
Female			-0.0354***	[0.0023]	-0.0203***	[0.0024]
Age >64			-0.0014	[0.0016]	-0.0016	[0.0018]
Age <45			-0.0057***	[0.0017]	-0.0023	[0.0017]
Low education			-0.0103***	[0.0025]	-0.0257***	[0.0027]
High education			0.0134***	[0.0023]	0.0075***	[0.0028]
Partner			-0.0022	[0.0020]	0.0006	[0.0027]
HH income: 1st quart.			-0.0033	[0.0021]	-0.0000	[0.0027]
HH income: 2nd quart.			-0.0027	[0.0019]	-0.0013	[0.0022]
HH income: 3rd quart.			-0.0020	[0.0015]	0.0003	[0.0018]
No. children in HH			0.0007	[0.0008]	-0.0005	[0.0008]
Constant	-0.0386***	[0.0015]	-0.0114***	[0.0030]	-0.0128***	[0.0037]
σ_1^*	0.1176***	[0.0007]	0.1185***	[0.0007]	0.1167***	[0.0007]
σ_2^*	0.5445***	[0.0084]	0.5533***	[0.0087]	0.5767***	[0.0095]
σ_3^*	0.2775***	[0.0038]	0.2791***	[0.0037]	0.2601***	[0.0032]
σ_{CDFfit}	0.1597***	[0.0004]	0.1602***	[0.0004]	0.1596***	[0.0004]
LogLik	-332,714.34		-331,725.45		-331,997.03	
AIC	665,500.68		663,630.90		664,172.05	
Observations	14,282		14,264		14,264	

Main equation: $\mu_{it}^* = \gamma_k R_{t-1} + \boldsymbol{x}_{it}\beta + \alpha_i + u_{it}$

	Full model	
γ_1: Return coeff. CI1	0.0286***	[0.0014]
γ_2: Return coeff. CI2	-0.5890***	[0.0166]
γ_3: Return coeff. CI3	0.7143***	[0.0127]
Female	-0.0354***	[0.0023]
Low education	-0.0103***	[0.0025]
High education	0.0134***	[0.0023]
Constant	-0.0114***	[0.0030]
Observations	14264	

- γ_1 is not restricted to zero in this specification
  - estimate of γ_1 is close to zero (corresponds to type RW)
- γ_2 and γ_3 are also unrestricted
  - estimate of γ_2 is negative (corresponds to type MR)
  - estimate of γ_3 is positive (corresponds to type P)
- Heterogeneity in the population as known from the literature

Type probabilities (RE multinomial logit)

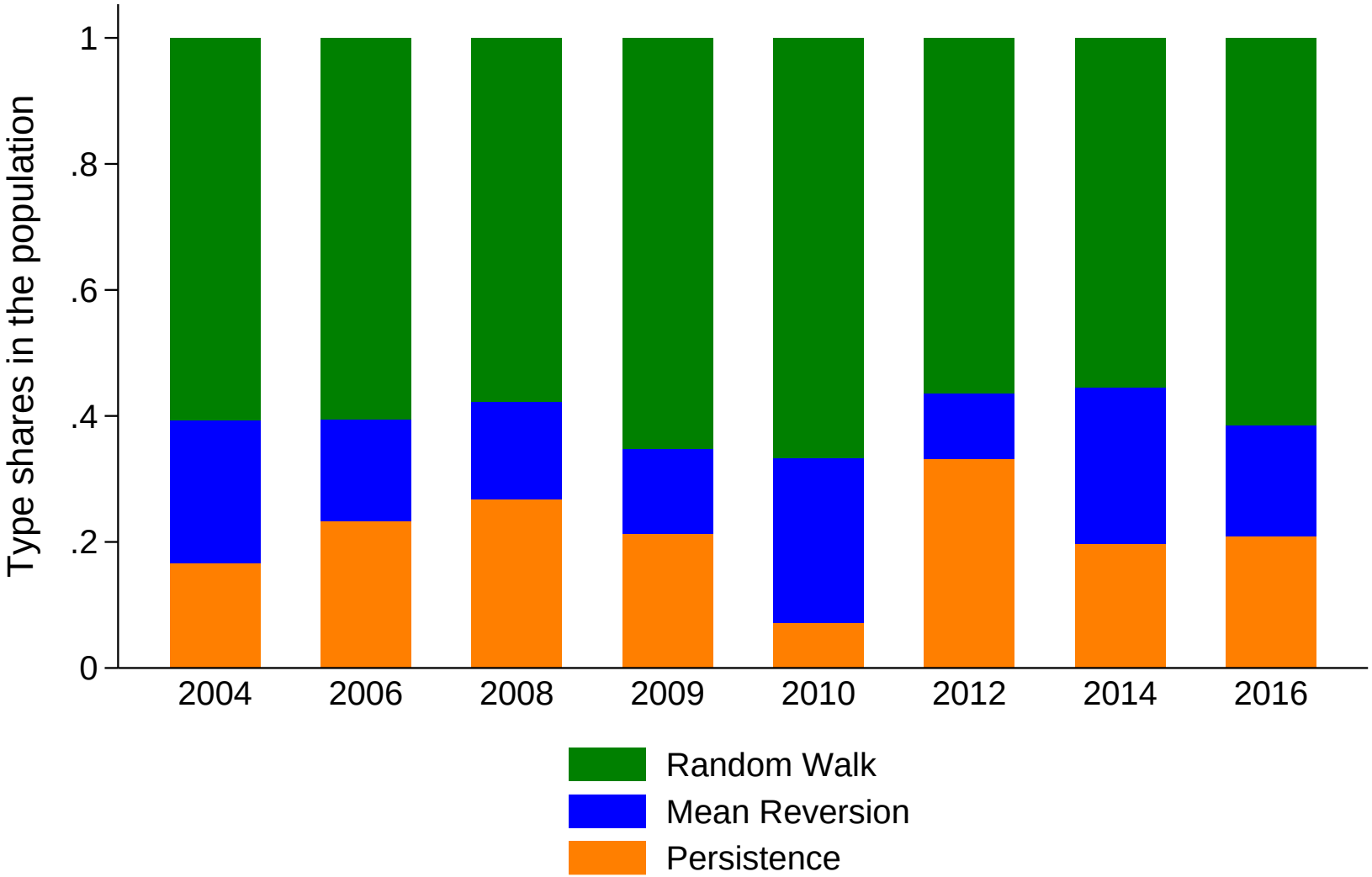
	Full model	
Class 2 (Mean Reversion)		
Female	0.6120***	[0.0713]
Age >64	-0.5586***	[0.0901]
Age <45	0.6899***	[0.0747]
Low education	0.0084	[0.0860]
High education	-0.6089***	[0.0816]
Partner	0.2717***	[0.0878]
HH income: 1st quart.	0.6510***	[0.1035]
HH income: 2nd quart.	0.3570***	[0.0942]
HH income: 3rd quart.	0.2764***	[0.0891]
No. children in HH	-0.0152	[0.0333]
Constant	-1.9978***	[0.1467]
Class 3 (Persistence)		
Female	0.4790***	[0.0709]
Age >64	-0.2835***	[0.0841]
Age <45	0.5440***	[0.0763]
Low education	0.0870	[0.0872]
High education	-0.3230***	[0.0811]
Partner	0.3157***	[0.0878]
HH income: 1st quart.	0.4740***	[0.1036]
HH income: 2nd quart.	0.2834***	[0.0934]
HH income: 3rd quart.	0.2743***	[0.0867]
No. children in HH	-0.0362	[0.0342]
Constant	-2.2920***	[0.1532]
Implied Cl1 share	0.62	
Implied Cl2 share	0.19	
Implied Cl3 share	0.19	
Observations	14264	

- Omitted category: Class 1 (RW)
- Heterogeneity in the population:
  - Females and young people are less likely to be in class 1 (RW)
  - Highly educated people are more likely to be in class 1 (RW)
- Implied type shares (RW,MR,P) are roughly (0.60,0.18,0.22)

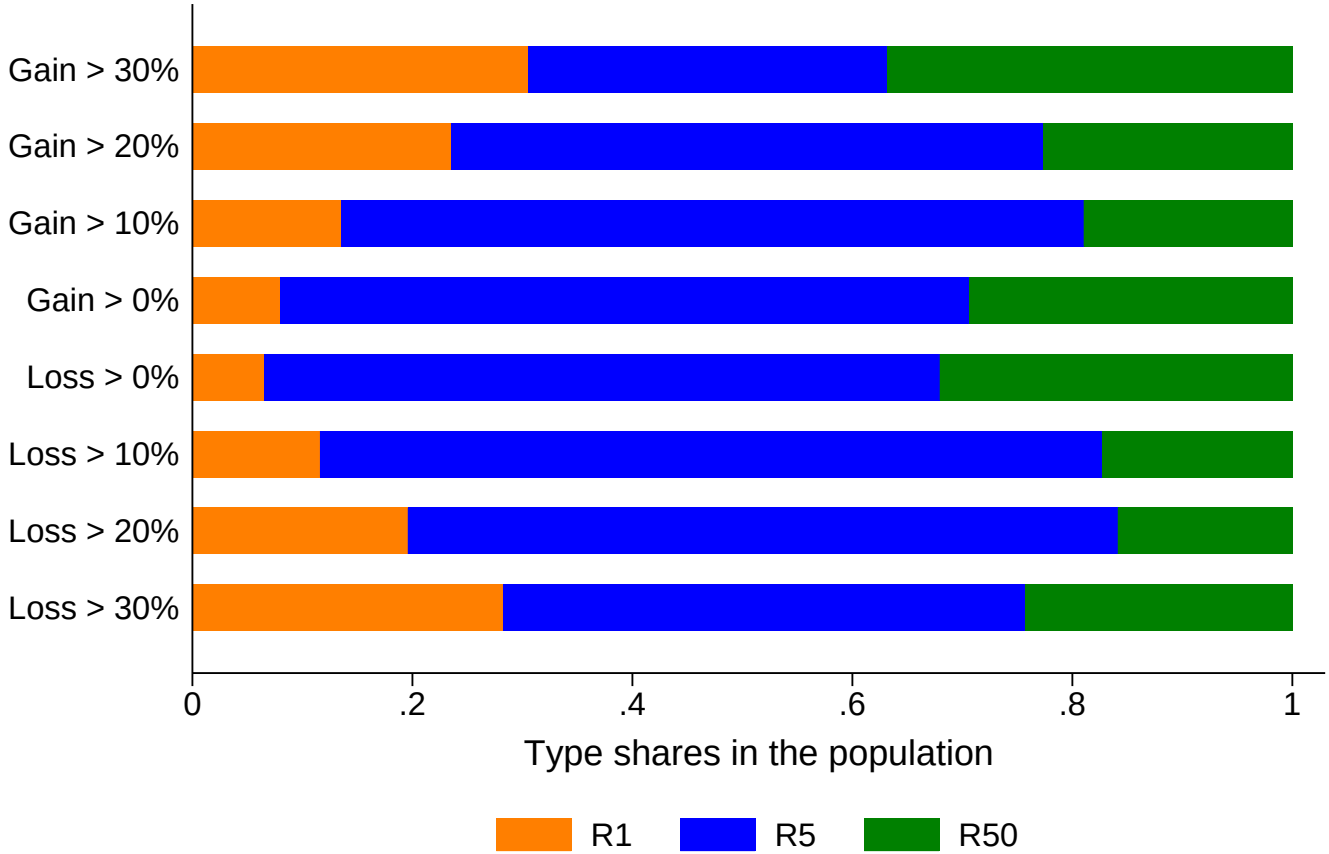
Type probabilities (RE multinomial logit)

	(1) Constants only		(2) Full model		(3) Restricted return coeff.	
Class 2 (Mean Reversion)						
Female			0.6120***	[0.0713]	0.3394***	[0.0715]
Age >64			-0.5586***	[0.0901]	-0.5734***	[0.0936]
Age <45			0.6899***	[0.0747]	0.7835***	[0.0766]
Low education			0.0084	[0.0860]	0.2229**	[0.0889]
High education			-0.6089***	[0.0816]	-0.5621***	[0.0870]
Partner			0.2717***	[0.0878]	0.1975**	[0.0944]
HH income: 1st quart.			0.6510***	[0.1035]	0.5987***	[0.1070]
HH income: 2nd quart.			0.3570***	[0.0942]	0.3336***	[0.0974]
HH income: 3rd quart.			0.2764***	[0.0891]	0.2305**	[0.0919]
No. children in HH			-0.0152	[0.0333]	0.0011	[0.0338]
Constant	-1.3143***	[0.0421]	-1.9978***	[0.1467]	-1.9628***	[0.1543]
Class 3 (Persistence)						
Female			0.4790***	[0.0709]	0.2216***	[0.0661]
Age >64			-0.2835***	[0.0841]	-0.2922***	[0.0809]
Age <45			0.5440***	[0.0763]	0.5227***	[0.0733]
Low education			0.0870	[0.0872]	0.2687***	[0.0834]
High education			-0.3230***	[0.0811]	-0.2241***	[0.0795]
Partner			0.3157***	[0.0878]	0.2552***	[0.0879]
HH income: 1st quart.			0.4740***	[0.1036]	0.3926***	[0.1008]
HH income: 2nd quart.			0.2834***	[0.0934]	0.2189**	[0.0900]
HH income: 3rd quart.			0.2743***	[0.0867]	0.2341***	[0.0833]
No. children in HH			-0.0362	[0.0342]	-0.0154	[0.0322]
Constant	-1.1492***	[0.0444]	-2.2920***	[0.1532]	-2.0169***	[0.1497]
Implied Cl1 share	0.59		0.62		0.60	
Implied Cl2 share	0.20		0.19		0.18	
Implied Cl3 share	0.21		0.19		0.21	
LogLik	-332,714.34		-331,725.45		-331,997.03	
AIC	665,500.68		663,630.90		664,172.05	
Observations	14,282		14,264		14,264	

Predicted type shares over time



Rounding shares (question-specific)



Robustness

- Unrestricted return coefficient: $\gamma_1 \neq 0$
- Short-run returns: 8,4,1 week(s) (rather than one year)
- Stock owners vs. non-owners
- Pre/post crisis

Summary

- Dynamic panel data model with a finite mixture of updating types
- Data are consistent with (RW,MR,P) types
- Distribution is roughly (.6,.18,.22)
- Probability of type membership varies across the **population** and over **time**
  - → Men and highly educated respondents are more likely to be RW
  - → Financial crisis increases shares of RW and MR types
  - → But only in the short run: Type distributions in 2014/16 similar to those in 2004/06