Online appendix to

"Does Money Illusion Matter?: Reply"

by Ernst Fehr and Jean-Robert Tyran

I. Does the design in Fehr and Tyran (2001) enable the measurement of individual-level money illusion?

Peterson and Winn (2013) claim that in our design we cannot measure individual-level money illusion: **"***In the treatment they* [FT] *used to measure individual-level money illusion rational agents and agents with money illusion have the same optimal strategy. Consequently, money illusion cannot be measured in their experiments*". In the following we show, that PW's failure to distinguish between behavioral money illusion and the various potential psychological mechanisms behind money illusion is the basis for this erroneous assertion. Their claim above refers to our treatments in which subjects played against 3 computerized opponents who were programmed to play a coordinated best reply to the subject's price choice. Subjects were informed about the aggregate reaction function of their computerized counterparts and were randomly assigned to play this game under a real payoff representation (RC) or under a nominal payoff representation (NC). We found that the subjects in NC converged *slightly* less quickly to the equilibrium of this game after a nominal shock and took this as evidence for a *small* amount of individual-level money illusion.

Subjects in this game basically faced an individual optimization problem in which the optimal solution requires taking the computerized players' aggregate response function into account. PW argue that *under the assumption that subjects always take the computerized opponents' aggregate best reply function fully into account* when choosing their strategy, the best strategy for a fully rational agent who maximizes her real payoff coincides with that of an agent who maximizes her nominal payoff. Indeed, if one imposes this restriction on the agents' play (i.e. forces them to take the aggregate best reply of their computerized opponents fully into account), the behavior of a rational agent and a nominal payoff maximizer would be identical in FT. In other words, PW's claim is only correct under their narrow definition of money illusion, *and* only under the additional assumption that subjects with this type of money illusion are otherwise rational and forward looking (i.e., taking the opponents' aggregate best

response fully into account). However, PW's claim above is a general assertion, and they fail to explicitly state the restrictive assumptions for the validity of their claim anywhere in their paper.

Moreover, there is no reason to restrict attention to PW's narrow psychological mechanism of money illusion. For example, why should an irrational agent who aims to maximize her nominal payoff not also make the mistake of at least partially neglecting the opponents' aggregate best reply function? It is perfectly possible that some subjects fail to always take the computerized agents' aggregate response function fully into account under the nominal representation because subjects have to deflate their nominal payoffs and the cognitive load of the nominal payoff representation is therefore higher. In fact, if subjects neglect the aggregate response function, they might believe that higher nominal payoffs are available (than those along the aggregate best response). There is thus absolutely no compelling reason to exclude these other forms of irrationality that may arise under a nominal frame which implies that FT can identify money illusion in these treatments.[1]

II. Best reply behavior

In Table A1 below, we document that subjects play a best reply to the expected average price regardless of whether they are in the real or the nominal treatment. Subjects' post-shock price choices are the dependent variable in all regressions of Table A1. Models (1) and (2) are based on price data after the negative monetary shock, models (3) and (4) on the post-shock data of the treatment with a positive monetary shock. The coefficient for the best reply to subjects' belief (BRB) about others' average price is close to one in all four models, and the hypothesis that the coefficient is equal to one is not rejected in any of the four cases (the 95% confidence interval includes the value "1" in all four models). In addition, the direct impact of the nominal treatment and the interaction between the best reply and the dummy variable for the nominal treatment is small and insignificant in all four models. Thus, the best reply to the expected average price determines almost one-for-one subjects' actual price choices, regardless of whether subjects are in the nominal or the real treatment.

[1] We thus agree with PW that our design does not enable us to identify *unconditional* nominal income maximization *at the individual level* as a source of money illusion. However, we never argued in our paper for this form of money illusion because – as explained in our reply – the hypothesis that subjects use nominal income as a proxy for real income is different. It requires that subjects display some other form of bounded rationality (e.g., cognitive load generated by the nominal frame) that makes them unable to pierce the veil of money.

We included the adaptive best reply (ABR, as defined by PW) into the regressions in Model (2) and Model (4). Interestingly, the coefficient of ABR and the interaction term between ABR and the nominal treatment are close to zero and insignificant, indicating that adaptive best reply behavior has no explanatory power once elicited expectations are controlled for. Moreover, the inclusion of the ABR and the interaction between NH and ABR leaves the explained variance in subjects' price choices (i.e. the R^2) completely unaffected. Table A1 therefore supports the conclusion that subjects play a best reply to their price expectations regardless of the treatment and the nature of the shock, and that the lagged average price does not affect behavior for given expectations.[2]

Do the results of Table A1 imply that money illusion does not affect subjects' price setting behavior? PW suggest this conclusion on the basis of a regression analysis similar to that in Table A1 (albeit without subjects' expectations, i.e., their BRB) because the treatment dummy NH and the interaction between NH and the ABR is not significant in their regressions. This conclusion is not valid, however, because it neglects the possibility that the nominal treatment affects subjects' expectations. We also find in the regressions in Table A1 that the nominal treatment does not affect subjects' price setting behavior if expectations are taken as given. If, however, the nominal treatment systematically affects subjects' expectations, money illusion has an indirect effect on subjects' price setting behavior via expectations. Therefore, the next question is how the treatment conditions and the lagged average price affected subjects' price expectations. We answer this question with Table 1 in the main text of our reply, which regresses subjects' expected average price on the deviation of the lagged average price from the post-shock equilibrium, the nominal treatment dummy, and the interaction of the two variables.

[2] These conclusions also hold if we neglect (as PW do) the data from the first post-shock period.

Table A1: Nominal price choices in post-shock periods in FT (2001)

	Negative shock		Positive shock	
Dep. Var.: Nominal Price	(1)	(2)	(3)	(4)
Best reply to belief (BRB) about others' average price	.936***	.914***	.984***	.981***
	(.037)	(.066)	(.014)	(.010)
Nominal treatment (NH = 1)	.218	.062	-.126	-.196
	(.376)	(.366)	(.589)	(.617)
BRB × NH	-.003	.025	.009	.048
	(.050)	(.099)	(.023)	(.053)
Best reply to lagged average price (ABR)		.020		.002
		(.030)		(.009)
ABR × NH		-.008		-.038
		(.070)		(.038)
Constant	.506**	.509**	.321	.322
	(.243)	(.249)	(.313)	(.305)
	R^2 = .842	R^2 = .842	R^2 = .897	R^2 = .897
	N = 1680	N = 1680	N = 1360	N = 1360

Notes: Random effects GLS regression, standard errors are in parentheses (clustered by group). Regressions are based on data from all post-shock observations in the treatments with human opponents (NH and RH) of FT (2001). Significance: * = 10 percent, ** = 5 percent, *** = 1 percent.

III. Evidence for money illusion at the individual level in FT and PW from post-experimental questionnaires

Section 4 in the Reply argues that the veil of money has a significant effect on behavior due to its effect on expectations. However, this evidence does not tell us whether (i) expectations are sticky in the NH because subjects themselves exhibit money illusion or (ii) because subjects believe that the other players have money illusion. We provide questionnaire evidence in FT that suggests that both reasons play a role. After the experiment, 49% of the subjects in the nominal treatment agreed with the statement "I interpreted high nominal incomes as an indicator for high real payoffs", and a majority of 53% of the subjects agreed with the statement "I believed that the other subjects would interpret high nominal payoffs as an indicator for high real payoffs." These data suggest that many subjects did in fact treat nominal payoffs as a proxy for real payoffs and also expected others do to so.

Moreover, important further data exist in PW that suggest the existence of individual-level money illusion. PW report questionnaire data in the unpublished discussion paper version (PW 2011) of their study that convey a very similar message as the data in FT mentioned above, but they omit these data in the AER version of their paper. In the discussion paper, they report that 52.3% of the subjects in the nominal treatment with a negative shock (60% in the nominal treatment with a positive shock) agree with the statement "I interpreted high nominal incomes as an indicator for high real payoffs". Likewise, 63.6 % of the subjects in the negative shock treatment (60% in the positive shock treatment) agree with the statement that "I believed that the other subjects would interpret high nominal payoffs as an indicator for high real payoffs." Thus, many subjects in PW's data also seem to take nominal income as a proxy for real income and believe that others do so – an observation that is consistent with the significant level of behavioral money illusion documented in Section 3 of the main paper (see also Figure 1).

IV. Does adaptive best reply behavior explain asymmetric price adjustment?

PW argue that the asymmetry in price adjustment across the negative and the positive shock is an artifact of the experimental design (i.e., the payoff structure) in FT. They argue that *if subjects play an adaptive best reply* (to the lagged average price), nominal prices *necessarily* adjust more quickly after a positive shock, i.e. they claim that the adjustment asymmetry exists regardless of the existence of

money illusion. The following facts cast doubts on this argument, however. First, as we show in Table A1, the adaptive best reply strategy has no predictive power for subjects' price choices once elicited expectations are controlled for. Rather than playing a best reply to the lagged average price, subjects predominantly play a best reply to their expectation of the opponents' current average price. Second, Table 1 in the paper shows that the nominal treatment has a significant impact on expectation formation, indicating that money illusion affects subjects' price expectations. Third, PW's own data (in which they control for differences in adjustment dynamics under an adaptive best reply strategy) show a pronounced asymmetry in price adjustment across shocks (see Figure 1). Apparently, there must be other forces than adaptive best reply dynamics at work that explain the asymmetry in price adjustments across shocks.