ONLINE APPENDIX for "Who wears the trousers in the family? Intra-household resource control, subjective expectations and human capital investment"

Validation of subjective expectations in Macedonia

The availability of credible measures for factors that enters the decision process (i.e. expectations) is fundamental to estimate structural models of individual behavior under weaker assumptions (Manski, 2004). Collecting accurate and credible data about subjective expectations turns then to be important in data collection activities that requires the gathering of this kind of information. While the collection of data about subjective expectations has been increasingly important in developed countries (see Manski, 2004), the topic gathered attention only recently in developing countries. This is mainly due to the fact that questions related to probabilities might not be fully understood or might require a higher burdensome in countries where the average level of education of respondents (and perhaps of enumerators) is low. In the other direction, data collected in developing countries are central since they might partly relate to the presence of market failures related to information and knowledge.

There is still no consensus in literature about the best practice to collect subjective expectations in developing countries. However, a growing and recent research agenda (an introduction about the strategies and the experiences about collecting expectations in developing countries is provided by Attanasio, 2009; Delavande et al., 2010) provides important insights on how to design and implement a strategy to collect information about subjective probabilities.

In literature, data about expectation have been mainly collected by using non-probabilistic methods (such as Likert scales) and more recently by using probabilistic methods with or without visual aids. Even if Likert scales have proven to be partially related to subjective probabilities (Delavande and Kohler, 2009), non-probabilistic method might be problematic for inter-personal comparisons since we can't ascertain which is the quantity reported by the respondent, while by eliciting subjective probabilities we might be able to recover the moments of a distribution of interest. However, eliciting subjective probabilities might be problematic in developing countries, since the concept of probability might not be understood by respondents with lower levels of education. For this reason, asking directly about probabilities (for instance, in the case of a developed country like Dominitz and Manski (1997), who ask about the probability of being employed in a future date in the US, and for a developing country like McKenzie et al. (2007), who ask about income expectations conditional on migration versus non-migration among Tongan migrants in New Zealand and Tongan residents) might not return credible results.

Table 1: Complete response rates for expectations related to schooling by gender of the child

	Baseline (2010)		Follow-up (2012)	
	Female	Male	Female	Male
Expectations for primary school	0.926	0.937	0.933	0.967
Expectations for secondary school	0.946	0.952	0.940	0.971
Expectations about employment	0.970	0.976	0.996	0.996
Probability to go to university	0.970	0.972	0.993	0.996

Note. An observation is considered complete if the respondent answers all requested information to compute expectations. Response rates are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). Response rates are divided by gender since some households report expectations for more than one child when children in the age range for completing the expectations section have different gender.

In this study, given the main characteristics of households in the sample, with high poverty and low levels of schooling, it is then important to select a methodology that would allow to elicit a credible measure of subjective expectations, while avoiding asking directly about probabilities. Following Attanasio et al. (2005), who collected income expectations in Colombia, andAttanasio and Kaufmann (2009), who elicited income expectations of junior high school students in Mexico, in the Republic of Macedonia, data about expectations on education returns from primary school and from secondary school have been collected using a method that elicit subjective expectations without directly using probabilistic terminology. Under both assumptions of having completed either primary or secondary school, it was asked about the minimum and the maximum of what the child could have earned, and after computing a mid-point, it was asked to report on a scale from 0 to 100 "how likely" the child would have earned less (or more) than the mid-point. The questionnaire asks to the household head (or its spouse in case of absence during the interview) information over the expected salary conditional on completion of primary or secondary school for at least one adolescent child in the household (in the case that are present two adolescents of different gender the information is collected for both).

The interesting feature is that data have been collected in a developing country, where evidence is still growing, and among its poorest quintiles (social financial assistance recipients), where respondents have a low level of schooling and have a lower experience in the formal labour market. In order to elicit subjective probabilities, it was used a visual aid strategy composed by a 0-100 ruler, that was initially presented with an example linking the probability of rain with the chosen scale (examples of more complex visual aids, such as the use of stones and coffee beens are provided by Luseno John and Winnie, 2003; Lybbert et al., 2007; Hill, 2006). The decision was linked to the fact that in the sample a large part of households live in rural areas and that rain is often a problem related to availability of utilities in urban areas.

Table 1 reports the response rates for the section about expectations. We can note that response rates are high and above 90% for all type of questions. Response rates are slightly higher for boys and for questions that involve a single answer. When facing more complex questions, such as the ones to elicit subjective expectations of the income distribution, response rates tend to be lower. Additionally response rates are slightly higher at follow-up compared to baseline, but the reasons are not clear (learning from the respondent, selection of the respondents or higher experience from the interviewers).

Testing the validity of subjective expectations

In order to understand whether collected answers are valid, it is important to answer the following question: does individuals understand probability questions? Are answers related to observable variables? How accurate are the answers?

In order to answer the first question, during the 2010 and 2012 data collection, we followed a strategy similar to Attanasio et al. (2005). In each municipality, households were randomly allocated into two groups: one group faced first the question about how likely is to earn less (or equal) than the calculated mid-point (we defined this group "X") and the other who faces first the question about how likely is to earn more than the calculated mid-point (group "Y"). We can then test whether the sum of the sample means for each group sums up to one. Tables 2 and 3 compare the means of the reported probability in the two groups and the sum of the means for the first answer and for the second answer reported by the respondent. As we can note the first answer perform much better than the second answer. When considering the first answer, for both boys and girls and for primary and secondary school expectations, we cannot reject the null hypothesis of the sum being different than 1 for most of the cases. The test perform much better for expectations collected at baseline. However, if we consider the second question answered, we reject the null hypothesis in all cases and the sum is always significantly lower than 1. This characteristics show that while the first reported answer conform to probability theory, the second answer is significantly lower than zero. One explanation is that, during the second questions, respondents tend to reduce their mental effort.

Unlike many applications, the questionnaire ask both questions to each respondent and allows then to test whether individual answers conform to probability theory by summing up to one. In other words, given the event A "earning an income between the minimum and the mid-point (included)", its complement $\bar{A}$ "earning an income between the mid-point and the maximum" and the event B "being employed", it is important to test whether the following condition is respected:

$$P(A \cup \bar{A}|B) = P(A|B) + P(\bar{A}|B) = P(A|B) + [1 - P(A|B)] = 1$$

Table 2: Testing for mean probability among groups to sum up to 1: first answer

	Baseline (2010)				Follow-up (2012)			
	Obs	Mean X	Mean Y	Sum	Obs	Mean X	Mean Y	Sum
Expectations after primary school								
Girls								
All observations	676	.504	.518	1.021	861	.456	.478	.935**
				(.024)				(.032)
Exclude 0 and 1	648	.506	.529	1.035	809	.455	.473	.928**
				(.023)				(.030)
Exclude .5, 0 and 1	436	.507	.527	1.033	497	.437	.455	.893**
				(.032)				(.042)
Boys								
All observations	718	.513	.538	1.051**	962	.487	.488	.976
				(.023)				(.029)
Exclude 0 and 1	704	.511	.540	1.052**	915	.483	.483	.965
				(.023)				(.028)
Exclude .5, 0 and 1	493	.514	.537	1.052	543	.462	.451	.913**
				(.032)				(.041)
Expectations after secondary school								
Girls								
All observations	674	.506	.530	1.036	863	.465	.477	.942**
				(.024)				(.027)
Exclude 0 and 1	646	.509	.534	1.044*	810	.470	.474	.945**
				(.023)				(.025)
Exclude .5, 0 and 1	434	.515	.548	1.063*	498	.451	.459	.910**
				(.033)				(.039)
Boys								
All observations	715	.502	.538	1.040*	962	.485	.488	.973
				(.023)				(.024)
Exclude 0 and 1	701	.502	.535	1.038	915	.479	.489	.968
				(.023)				(.023)
Exclude .5, 0 and 1	489	.504	.550	1.053	543	.465	.482	.947
				(.033)				(.038)

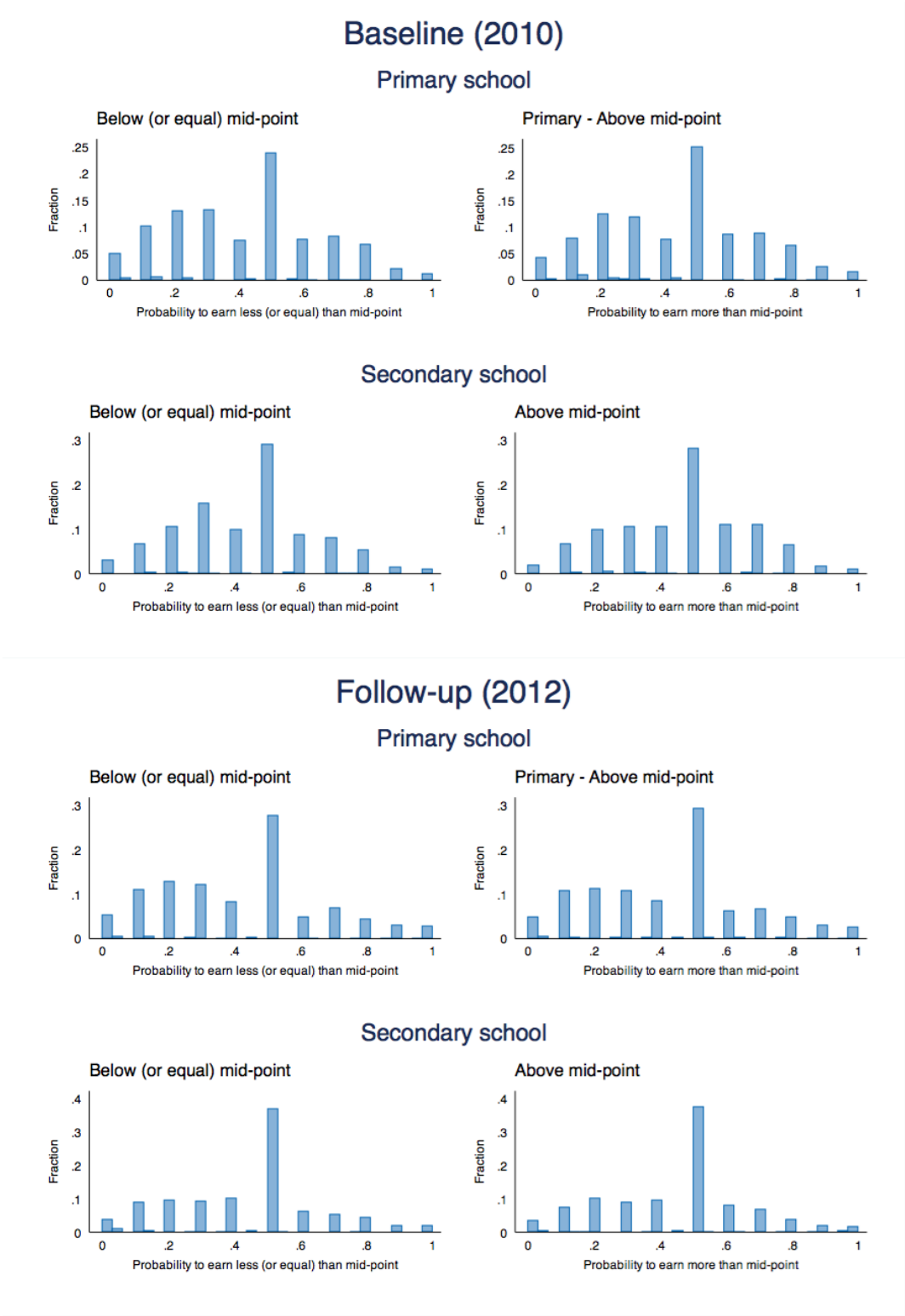
Note. In parenthesis, I report the share of total observation in the category. Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). Expectations are divided by gender since some households report expectations for more than one child when children in the age range for completing the expectations section have different gender. Standard errors are clustered at municipality level and * 0.10 ** 0.05 *** 0.01 represents the statistical significance of a test of equality to one of the sum of probabilities.

Table 3: Testing for mean probability among groups to sum up to 1: second answer

	Baseline (2010)				Follow-up (2012)			
	Obs	Mean X	Mean Y	Sum	Obs	Mean X	Mean Y	Sum
Expectations after primary school								
Girls								
All observations	676	.354	.309	.663***	861	.372	.346	.718***
				(.024)				(.020)
Exclude 0 and 1	644	.359	.323	.682***	813	.385	.350	.735***
				(.024)				(.020)
Exclude .5, 0 and 1	463	.333	.299	.633***	494	.360	.315	.675***
				(.027)				(.024)
Boys								
All observations	718	.349	.331	.680***	962	.365	.349	.714***
				(.025)				(.021)
Exclude 0 and 1	696	.352	.338	.689***	921	.371	.355	.726***
				(.025)				(.021)
Exclude .5, 0 and 1	509	.331	.315	.647***	565	.341	.328	.669***
				(.028)				(.028)
Expectations after secondary school								
Girls								
All observations	674	.391	.347	.738***	863	.399	.378	.777***
				(.023)				(.019)
Exclude 0 and 1	642	.399	.366	.765***	815	.419	.389	.808***
				(.021)				(.020)
Exclude .5, 0 and 1	461	.356	.319	.674***	496	.363	.323	.686***
				(.025)				(.025)
Boys								
All observations	715	.407	.369	.776***	962	.407	.383	.790***
				(.023)				(.021)
Exclude 0 and 1	693	.413	.379	.792***	920	.417	.396	.813***
				(.022)				(.020)
Exclude .5, 0 and 1	505	.378	.339	.716***	564	.366	.328	.694***
				(.026)				(.024)

Note. In parenthesis, I report the share of total observation in the category. Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). Expectations are divided by gender since some households report expectations for more than one child when children in the age range for completing the expectations section have different gender. Standard errors are clustered at municipality level and * 0.10 ** 0.05 *** 0.01 represents the statistical significance of a test of equality to one of the sum of probabilities.

Figure 1: Distribution of expected income after under different distributional assumptions



Note. Panels titled "Primary school" presents probabilities related to the completion of primary school only, while panels titled "Secondary school" presents probabilities related to the completion of secondary school.

Table 4: Distribution and shares for the sum of reported probabilities

	Baseline (2010)		Follow-up (2012)	
	Female	Male	Female	Male
Expectations after primary school				
Sum smaller than 1	314	302	366	376
	(0.464)	(0.421)	(0.423)	(0.389)
Sum equal to to 1	297	349	444	512
	(0.439)	(0.486)	(0.513)	(0.530)
Sum larger than 1	65	67	55	78
	(0.096)	(0.093)	(0.064)	(0.081)
Expectations after secondary school				
Sum smaller than 1	255	268	312	321
	(0.378)	(0.375)	(0.360)	(0.332)
Sum equal to to 1	331	355	488	569
	(0.491)	(0.497)	(0.563)	(0.589)
Sum larger than 1	88	92	67	76
	(0.131)	(0.129)	(0.077)	(0.079)

Note. In parenthesis, I report the share of total observation in the category. Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). Expectations are divided by gender since some households report expectations for more than one child when children in the age range for completing the expectations section have different gender.

Figure **1** shows the distribution of answers for event A (left panel) and for its complement $\bar{A}$ (right panels). In order to test whether answers conform to probability law, I first computed the sum of reported probabilities ($P(A|B) + P(\bar{A}|B)$) and then tested whether they differ from one[1]. Table 4 presents the distribution of the sum of reported probabilities for both primary school and secondary school expectations. We can observe that the cases in which probabilities sum up to one are ranging across the two data collection waves from a minimum of 44 percent to a maximum of 59 percent, depending on the precise category for which expectations are collected (achieved educational level and gender of the child). There is a strong tendency from respondents to report sums that are smaller than one, with roughly only 10 percent of observations being larger than 1.

In order to understand how strong is the tendency of report sums smaller or larger than one, I look at the mean sum of probabilities for the observations that don't conform to probability law and compare them to the mean for all observations (Table 5). We can observe that the sum ranges from a minimum of 0.61 (0.56) to a maximum of 0.70 (0.60) for the baseline (follow-up) in the cases in which the sum is smaller than one. If we consider instead the few cases in which the sum is larger than one, means ranges from 1.16 (1.17) to 1.21 (1.22) for the baseline (follow-up). We find that in all cases, even if roughly half of the respondents provides answers that sum up to one, the mean probability is always

[1]Mahajan et al. (2008) run a similar test by collecting subjective expectations about income in India using 10 stones for the questions about likelihood to earn less than the mid-point and other 10 stones to indicate the likelihood to earn more than the mid-point; they find that only 513 out of 1945 individuals answers with probabilities summing up to one, but the mean sum of probabilities is equal to 1.13 (including all observations) or 1.06 (excluding answers equal to 0 and 1), that is encouraging given the fact that answers were limited to multiples of 0.1.

Table 5: Average sum of reported probabilities

	Baseline (2010)		Follow-up (2012)	
	Female	Male	Female	Male
Expectations after primary school				
Sum smaller than 1	0.618	0.641	0.560	0.566
	(0.274)	(0.254)	(0.268)	(0.260)
Sum larger than 1	1.208	1.175	1.200	1.172
	(0.152)	(0.100)	(0.139)	(0.120)
All observations	0.843	0.865	0.826	0.845
	(0.291)	(0.259)	(0.293)	(0.281)
Expectations after secondary school				
Sum smaller than 1	0.644	0.697	0.565	0.593
	(0.248)	(0.222)	(0.272)	(0.259)
Sum larger than 1	1.168	1.170	1.209	1.216
	(0.128)	(0.099)	(0.140)	(0.152)
All observations	0.887	0.908	0.860	0.882
	(0.254)	(0.222)	(0.283)	(0.262)

Note. Standard deviations in parenthesis. Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). Expectations are divided by gender since some households report expectations for more than one child when children in the age range for completing the expectations section have different gender.

significantly different than one, suggesting that it might be reasonable to use different specifications to construct the probability distribution, such as using rescaling of probabilities or using only the first answer. In Section "First answer versus rescaling" I discuss how these two methods interact with distributional assumptions, which are first introduced in the following section.

Distributional assumptions

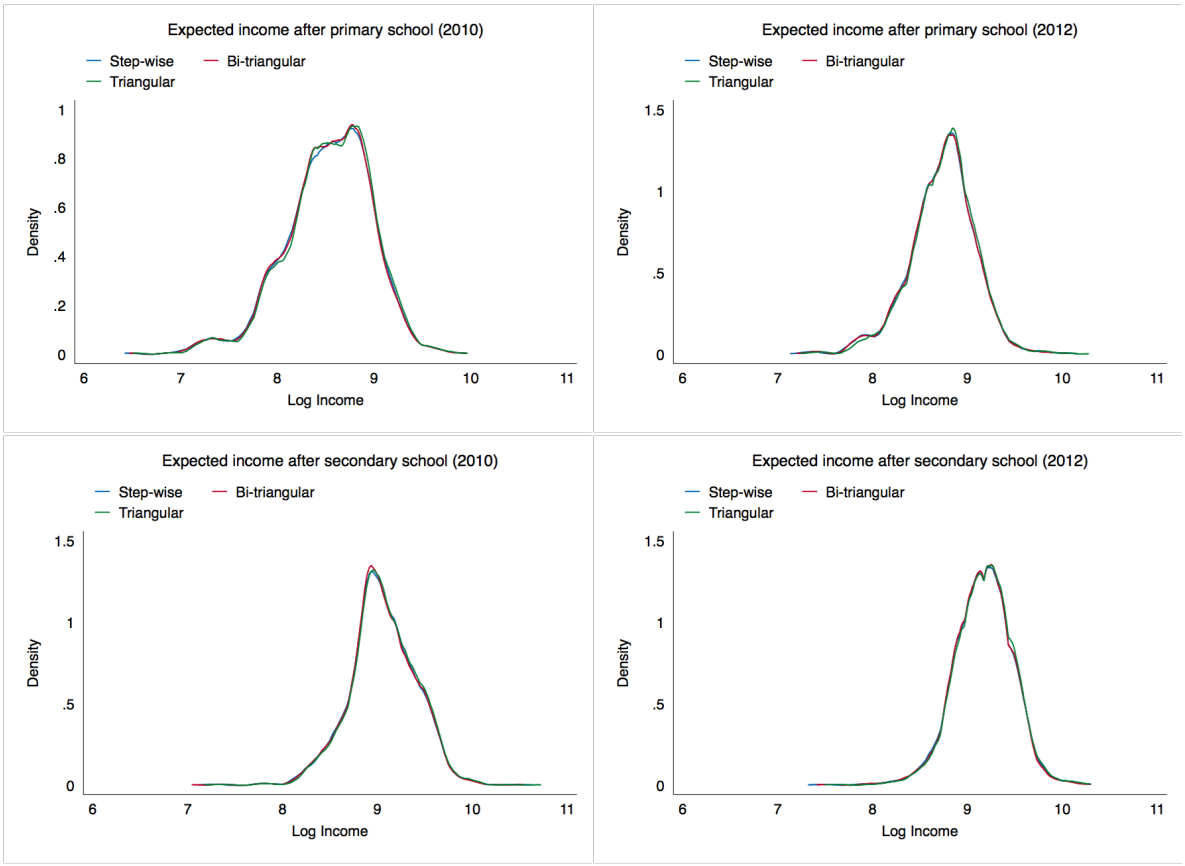
Given the structure of the collected information, assuming a specific class of distribution functions allows constructing the distribution of the expected salary and calculate its first moments. Specifically, assuming that a is the reported salary in the worst case, b is the reported salary in the best case and $f_Y(y)$ is the assumed continuous density function of the expected salary for one respondent, we can calculate the expected value and the variance using standard statistical formulas:

$$E[Y] = \int_a^b y\, f_Y(y)\, \mathrm{d}y = \mu \tag{1}$$

$$Var[Y] = \int_a^b (y-\mu)^2\, f_Y(y)\, \mathrm{d}y \tag{2}$$

In order to reconstruct the probability density function, it is necessary to consider distribution that can be identified using the a, b and the reported mass probability between a and the midpoint $(a+b)/2$. Distribution functions that are consistent with this setting are the step-wise uniform distribution, the triangular distribution and the bi-triangular distribution.

Figure 2: Distribution of expected income after under different distributional assumptions

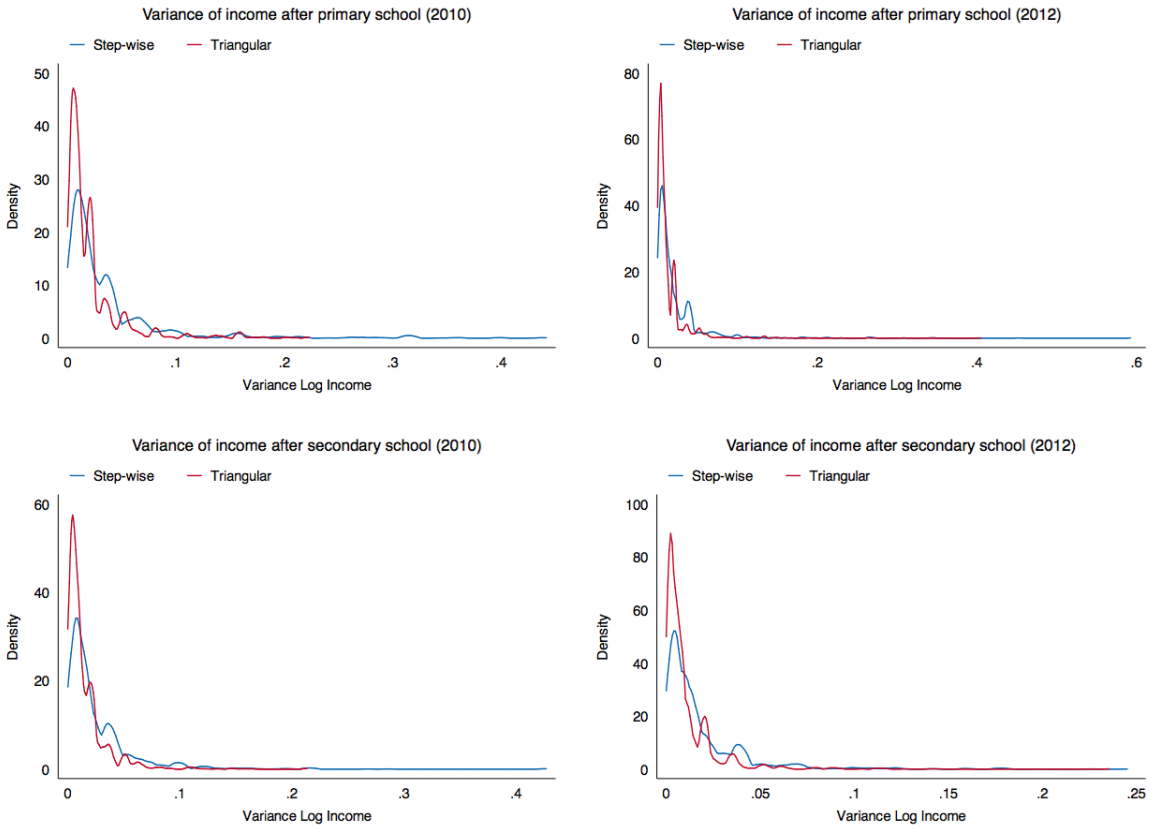


Note. Expected income is computed using log-income and using different distributional assumptions (step-wise, bi-triangular and triangular). The two top panels show the expected income after completion of primary school only at baseline and at follow-up, while the lower panels show the expected income after completion of secondary school at baseline and follow-up.

Figure 2 reports the sample distribution of the expected income conditional on completing primary and secondary school at baseline and follow-up using different distributional assumptions. We can observe that there is no strong difference between distributional assumptions for what concerns the distribution of the expected income, both conditional on completing primary and secondary school.

If we focus on the second moment of the distributions, we can note that differences become more significant. Figure 2 reports the sample distribution of the variance of income conditional on completing primary and secondary school at baseline and follow-up using different distributional assumptions (for clarity of the graph, I present only the distribution using only the step-wise uniform and the triangular distribution). An important characteristic to be researched among distribution functions that can be assumed for the purpose of constructing the distribution of subjective income expectations is that the density is decreasing while moving towards the extremes. This is the case that led the triangular distribution to be used in all the application of the paper. For this reason, we clearly expect the

Figure 3: Distribution of variance of income under different distributional assumptions



Note. Variance of income is computed using log-income and using different distributional assumptions (step-wise and triangular). The two top panels show the variance of income after completion of primary school only at baseline and at follow-up, while the lower panels show the variance of income after completion of secondary school at baseline and follow-up.

distribution of variances to differ between the two distributional assumptions.

First answer versus rescaling

As showed in Section , when the respondent is asked about the probability to earn below (or equal) and above a certain threshold, there is a possibility that the respondent reduce its mental effort during the second answer, such that the reported probabilities do not sum up to one. It is therefore important to understand whether using only the first answer or use both answers would lead to fundamental differences to the subjective income distribution.

In order to test for differences, I use compare the effect of choosing rescaling versus using the first answer only by looking at the distribution of expected income and its variance. Assuming that each respondent answers both questions about earning below (p_A) and above (p_B) the mid-point, the first strategy is to consider only the first question answered by the respondent ($\bar{p} = p_A$) and then compute the complement $(1 - \bar{p})$ as $1 - p_A$. The second strategy is to use both answers and rescale them in case the sum $p_A + p_B$ is different than one. In this case, $\bar{p}$ is simply determined by $p_A/(p_A+p_B)$.

Table 6 reports expected income for boys and gilrs at baseline and follow-up computed under different distributional assumption and using either a rescaled probability or a first answer probability. If we look at the mean difference between these two methods, we can note that differences are never significantly different than zero at baseline and rarely significant at follow-up. This provides evidence that using different methods do not lead to significant differences in computed expected incomes, even when considering different distributional assumptions. Additionally, we can compare the distibutions of expected income and its variance for each different method. Figure 4 and Figure 5 show that there are no significant differences even when we consider the whole sample distribution.

Reported income bounds, returns and individual characteristics

Another test that could be done in order to verify the validity of expectations is to control the relationship between the answers about minimum and maximum reported income with observable characteristics of the household and the respondent.

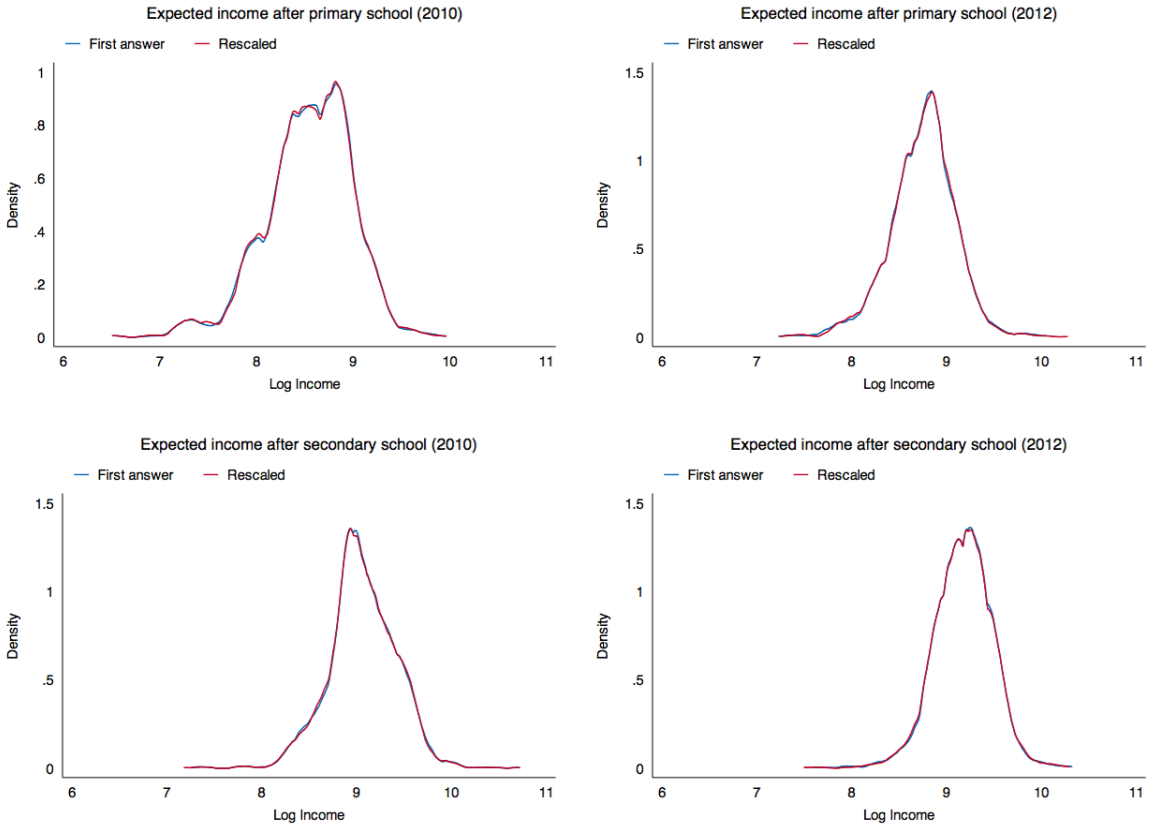
Table 7 and Table 8 present linear regressions of the minimum and maximum (log-)income reported by the respondent upon primary and secondary school completion and its difference (defined here as delta) on a series of individual and household characteristics. Expected income for male chidren are significantly higher in both rounds of data collection and for both primary and secondary school outcomes, but there is no significant difference across female and male children when considering the difference. At the same time, education of the household head has a role in explaining reported

Table 6: Comparison of expected income with first answer versus rescaling under different distributional assumptions

	Step-wise uniform		*Bi-triangular*		*Triangular*	
	Female	Male	Female	Male	Female	Male
Baseline (2010)						
Expectations after primary school						
Rescaled	8.482	8.541	8.480	8.541	8.502	8.563
	[0.448]	[0.464]	[0.444]	[0.459]	[0.444]	[0.456]
First answer	8.478	8.541	8.477	8.541	8.498	8.561
	[0.450]	[0.467]	[0.446]	[0.461]	[0.446]	[0.459]
Difference	0.002	0.000	0.001	0.000	0.003	0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Expectations after secondary school						
Rescaled	8.993	9.072	8.991	9.071	9.007	9.085
	[0.349]	[0.356]	[0.346]	[0.352]	[0.346]	[0.352]
First answer	8.992	9.070	8.990	9.070	9.006	9.084
	[0.348]	[0.355]	[0.346]	[0.351]	[0.347]	[0.352]
Difference	0.001	0.002	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Follow-up (2012)						
Expectations after primary school						
Rescaled	8.681	8.762	8.681	8.761	8.698	8.778
	[0.336]	[0.358]	[0.332]	[0.356]	[0.329]	[0.352]
First answer	8.680	8.762	8.679	8.761	8.699	8.780
	[0.338]	[0.358]	[0.333]	[0.355]	[0.328]	[0.350]
Difference	0.001	-0.000	0.000	-0.000	-0.002*	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Expectations after secondary school						
Rescaled	9.122	9.191	9.121	9.190	9.135	9.204
	[0.301]	[0.312]	[0.298]	[0.307]	[0.297]	[0.309]
First answer	9.118	9.189	9.118	9.188	9.133	9.202
	[0.304]	[0.317]	[0.300]	[0.311]	[0.297]	[0.311]
Difference	0.003*	0.002	0.002*	0.001	0.001	0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)

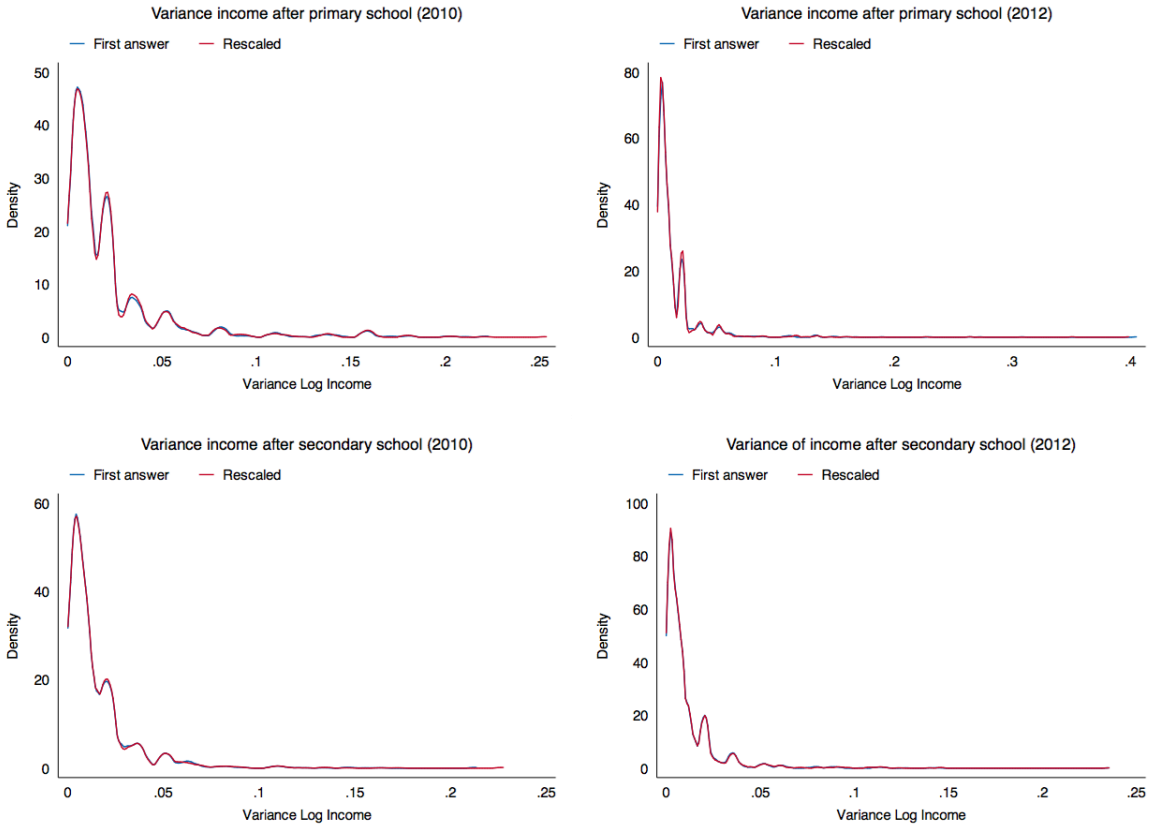
Note. Standard deviations in brackets, standard errors in parenthesis. "Rescaled" expectations are computed by rescaling the sum of reported probabilities to be equal to one, while "first answer" expectations are computed using the first reported answer only. The difference is computed for each observation as the difference between the "rescaled" and the "first answer" values. Standard errors are clustered at municipality level and * 0.10 ** 0.05 *** 0.01 represents the statistical significance of a t-test for equality to zero of the difference. Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). Expectations are divided by gender since some households report expectations for more than one child when children in the age range for completing the expectations section have different gender.

Figure 4: Distribution of expected income comparing first answer and rescaling



Note. Expected income is computed using log-income. "Rescaled" expectations are computed by rescaling the sum of reported probabilities to be equal to one, while "first answer" expectations are computed using the first reported answer only. Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). The two top panels show the expected income after completion of primary school only at baseline and at follow-up, while the lower panels show the expected income after completion of secondary school at baseline and follow-up.

Figure 5: Distribution of variance of income comparing first answer and rescaling



Note. Variance of income is computed using log-income. "Rescaled" expectations are computed by rescaling the sum of reported probabilities to be equal to one, while "first answer" expectations are computed using the first reported answer only. Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). The two top panels show the variance of income after completion of primary school only at baseline and at follow-up, while the lower panels show the variance of income after completion of secondary school at baseline and follow-up.

Table 7: Maximum and minimum income and individual characteristic (Baseline 2010)

	Primary school			Secondary school		
	Min	Max	Delta	Min	Max	Delta
Male child	0.054**	0.068***	0.002	0.093***	0.090***	-0.001
	(0.024)	(0.025)	(0.018)	(0.019)	(0.020)	(0.013)
Child age (years)	0.006	-0.000	-0.005	0.015***	0.005	-0.009***
	(0.007)	(0.006)	(0.004)	(0.005)	(0.005)	(0.003)
Male (head)	0.019	0.035	0.010	0.047	-0.038	-0.062
	(0.069)	(0.080)	(0.044)	(0.050)	(0.059)	(0.040)
Age head (less than 40 y.o.)	-0.013	0.007	0.019	0.040	0.005	-0.032
	(0.045)	(0.031)	(0.034)	(0.034)	(0.025)	(0.025)
Lower primary (head)	-0.024	-0.035	-0.004	-0.028	-0.079**	-0.037
	(0.053)	(0.042)	(0.030)	(0.034)	(0.033)	(0.025)
Secondary school (head)	-0.103**	-0.079**	0.021	-0.048	-0.052*	-0.001
	(0.043)	(0.040)	(0.027)	(0.030)	(0.029)	(0.020)
Albanian	0.111	-0.006	-0.112***	0.009	-0.087	-0.096***
	(0.079)	(0.066)	(0.039)	(0.055)	(0.059)	(0.030)
Roma	0.046	-0.021	-0.086*	0.001	-0.092*	-0.087**
	(0.063)	(0.054)	(0.045)	(0.048)	(0.048)	(0.037)
Turk	0.011	-0.024	-0.087*	0.064	0.048	-0.035
	(0.074)	(0.056)	(0.045)	(0.054)	(0.049)	(0.038)
Household members	-0.002	0.009	0.009	0.002	0.018	0.016*
	(0.023)	(0.013)	(0.016)	(0.015)	(0.013)	(0.009)
Boys 13-18 y.o.	0.039	-0.005	-0.052**	-0.005	-0.025	-0.026
	(0.032)	(0.029)	(0.023)	(0.021)	(0.022)	(0.016)
Girls 13-18 y.o.	0.026	0.011	-0.041**	0.009	-0.002	-0.019
	(0.034)	(0.031)	(0.018)	(0.023)	(0.024)	(0.015)
Father is present	-0.131	-0.091	0.075	-0.134*	-0.011	0.111**
	(0.085)	(0.083)	(0.050)	(0.072)	(0.068)	(0.050)
Mother is present	0.006	0.029	0.046	0.062	0.099	0.045
	(0.093)	(0.061)	(0.057)	(0.065)	(0.061)	(0.044)
Rural	0.070	-0.035	-0.112***	0.040	-0.052	-0.091***
	(0.054)	(0.045)	(0.032)	(0.049)	(0.047)	(0.030)
Part of City of Skopje	0.043	-0.100	-0.135	0.018	-0.082	-0.074
	(0.155)	(0.143)	(0.108)	(0.098)	(0.103)	(0.081)
Wealth (low)	-0.113**	-0.077**	0.025	-0.047	-0.037	0.012
	(0.048)	(0.034)	(0.038)	(0.033)	(0.031)	(0.025)
Wealth (high)	0.044	-0.027	-0.068*	0.055*	0.006	-0.053**
	(0.054)	(0.046)	(0.035)	(0.033)	(0.034)	(0.026)
Time in Fin.Ass. (1-6 years)	-0.017	0.032	0.048	-0.064*	-0.039	0.022
	(0.053)	(0.047)	(0.035)	(0.037)	(0.037)	(0.023)
Time in Fin.Ass. (> 6 years)	-0.084	-0.038	0.035	-0.065*	-0.059*	0.001
	(0.054)	(0.048)	(0.034)	(0.036)	(0.035)	(0.025)
Unemployment (<=0.30)	0.220**	0.237***	0.029	0.188**	0.265***	0.089*
	(0.109)	(0.082)	(0.058)	(0.072)	(0.068)	(0.047)
Constant	8.239***	8.790***	0.554***	8.630***	9.160***	0.515***
	(0.199)	(0.153)	(0.130)	(0.124)	(0.127)	(0.105)
Observations	1328	1354	1328	1364	1367	1364

Note. Standard errors in parenthesis (* 0.10 ** 0.05 *** 0.01). Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). Minimum and maximum income are reported in logarithms. Unemployment is computed at regional level for the year before the interview and ranges from 14.4 percent to 64.8 percent in 2009 and from 9.3 percent to 42.8 percent in 2011. Months in Financial Assistance are reported by the respondent. Omitted categories include: Female child, Macedonian and other ethnicities, Female head, Time in Financial Assistance (less than 1 year), Unemployment (larger than 30 percent). All specifications include regional dummies.

Table 8: Maximum and minimum income and individual characteristic (Follow-up 2012)

	Primary school			Secondary school		
	Min	Max	Delta	Min	Max	Delta
Male child	0.078***	0.080***	0.000	0.078***	0.077***	-0.002
	(0.020)	(0.017)	(0.012)	(0.017)	(0.014)	(0.010)
Child age (years)	0.006	0.005*	0.000	0.001	0.003	0.002
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)
Male (head)	0.043	0.034	-0.011	0.069*	0.010	-0.057*
	(0.050)	(0.040)	(0.039)	(0.039)	(0.027)	(0.034)
Age head (less than 40 y.o.)	-0.066	0.045	0.102**	0.056*	0.059*	0.003
	(0.055)	(0.034)	(0.043)	(0.033)	(0.031)	(0.026)
Lower primary (head)	-0.161***	-0.066**	0.097***	-0.070**	-0.073**	-0.005
	(0.045)	(0.031)	(0.035)	(0.032)	(0.031)	(0.023)
Secondary school (head)	0.035	0.048*	0.006	0.060**	0.061**	-0.000
	(0.032)	(0.025)	(0.023)	(0.025)	(0.024)	(0.029)
Albanian	0.047	0.031	-0.034	-0.005	-0.003	0.006
	(0.051)	(0.041)	(0.042)	(0.036)	(0.037)	(0.044)
Roma	0.065	0.003	-0.070*	0.004	-0.031	-0.035
	(0.058)	(0.041)	(0.039)	(0.039)	(0.040)	(0.027)
Turk	0.064	-0.028	-0.091**	0.022	-0.038	-0.056*
	(0.061)	(0.043)	(0.041)	(0.041)	(0.034)	(0.031)
Household members	0.004	0.003	-0.004	0.000	0.002	0.001
	(0.014)	(0.011)	(0.009)	(0.011)	(0.011)	(0.007)
Boys 13-18 y.o.	0.013	-0.005	-0.015	-0.003	-0.018	-0.014
	(0.024)	(0.019)	(0.015)	(0.018)	(0.017)	(0.014)
Girls 13-18 y.o.	-0.011	-0.002	0.009	-0.025	-0.007	0.019
	(0.023)	(0.017)	(0.016)	(0.017)	(0.016)	(0.013)
Father is present	-0.161**	-0.046	0.125***	-0.064	-0.019	0.042
	(0.061)	(0.051)	(0.047)	(0.050)	(0.039)	(0.046)
Mother is present	-0.030	-0.036	-0.003	-0.075	-0.036	0.039
	(0.072)	(0.059)	(0.063)	(0.076)	(0.070)	(0.035)
Rural	-0.028	-0.056	-0.031	-0.071	-0.088*	-0.018
	(0.055)	(0.042)	(0.043)	(0.044)	(0.045)	(0.038)
Part of City of Skopje	0.063	-0.022	-0.058	-0.044	-0.033	-0.000
	(0.112)	(0.107)	(0.080)	(0.067)	(0.061)	(0.060)
Wealth (low)	0.025	-0.030	-0.042*	0.020	0.016	-0.002
	(0.038)	(0.031)	(0.025)	(0.029)	(0.027)	(0.022)
Wealth (high)	0.073**	0.043	-0.031	0.017	0.037	0.021
	(0.031)	(0.030)	(0.029)	(0.030)	(0.030)	(0.030)
Time in Fin.Ass. (1-6 years)	-0.045	0.017	0.060**	-0.015	-0.012	0.002
	(0.038)	(0.032)	(0.029)	(0.028)	(0.029)	(0.023)
Time in Fin.Ass. (> 6 years)	-0.096***	-0.072**	0.029	-0.075***	-0.090***	-0.016
	(0.035)	(0.029)	(0.028)	(0.026)	(0.025)	(0.026)
Unemployment (<=0.30)	-0.026	0.115*	0.039	0.160***	-0.032	-0.123***
	(0.120)	(0.060)	(0.093)	(0.053)	(0.069)	(0.046)
Constant	8.431***	8.799***	0.364***	8.923***	9.299***	0.380***
	(0.141)	(0.097)	(0.111)	(0.123)	(0.104)	(0.088)
Observations	1415	1434	1415	1437	1442	1437

Note. Standard errors in parenthesis (* 0.10 ** 0.05 *** 0.01). Expectations are restricted to recipients of Social Financial Assistance and include all respondents (including resampled households at follow-up). Minimum and maximum income are reported in logarithms. Unemployment is computed at regional level for the year before the interview and ranges from 14.4 percent to 64.8 percent in 2009 and from 9.3 percent to 42.8 percent in 2011. Months in Financial Assistance are reported by the respondent. Omitted categories include: Female child, Macedonian and other ethnicities, Female head, Time in Financial Assistance (less than 1 year), Unemployment (larger than 30 percent). All specifications include regional dummies.

income, but with a different pattern at baseline and follow-up. While at baseline higher education lead to a lower expected income after primary school, af follow-up lower education is correlated with lower incomes while higher education of the household head is correlated with higher incomes. To provide evidence that reported expectations are linked to monetary returns in the labour market, I control for unemployment rate at regional level, by dividing munipalities into (relatively) low and high unemployment. We can note that larger unemployment in the year before the interview affect negatively reported incomes for both educational level, providing evidence that in areas with high unemployment respondents expect lower incomes.

References

Attanasio, O. (2009). Expectations and Perceptions in Developing Countries: Their Measurement and Their Use. *American Economic Review*, 99(2):87–92.

Attanasio, O. and Kaufmann, K. (2009). Educational choices and subjective expectations of returns: Evidence on intrahousehold decisions and gender differences. *NBER Working Paper 15087*.

Attanasio, O., Meghir, C., and Vera-Hernandez, M. (2005). Elicitation, validation, and use of probability distributions of future income in developing countries. *Paper prepared for the 2005 Econometric Society Meeting*.

Delavande, A., Giné, X., and McKenzie, D. (2010). Measuring subjective expectations in developing countries: a critical review and new evidence. *Journal of Development Economics*.

Delavande, A. and Kohler, H. (2009). Subjective expectations in the context of hiv/aids in malawi. *Demographic research*, 20:817.

Dominitz, J. and Manski, C. (1997). Using expectation data to study subjective income expectations. *Journal of the American Statistical Association*, 27 (1):30–41.

Hill, R. (2006). Coffee price risk in the market: Exporter, producer and trader data from uganda.

Luseno John, G. and Winnie, K. (2003). Assessing the value of climate forecast information for pastoralists: Evidence from southern ethiopia and northern kenya. *World Development*, 31(9):1477–1494.

Lybbert, T., Barrett, C., McPeak, J., and Luseno, W. (2007). Bayesian herders: Updating of rainfall beliefs in response to external forecasts. *World Development*, 35(3):480–497.

Mahajan, A., Tarozzi, A., Yoong, J., and Blackburn, B. (2008). Bednets, information and malaria in orissa. In *World Bank Poverty and Applied Micro Seminar Series, World Bank, Washington, DC (December 10)*.

Manski, C. (2004). Measuring expectations. *Econometrica*, pages 1329–1376.

McKenzie, D., Gibson, J., and Stillman, S. (2007). A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad? *World Bank Policy Research*, (Working Paper 4141).