A Appendix - for online publication

A.1 Figures

Figure 1: Location of Karonga District within Malawi

A.2 Additional data and results

Table 13: Individual economic activity & schooling – 2013 member census

Category	N	% of responses
Occupation	**3780**	**100%**
Farmer	2,786	73.34%
Business	793	20.87%
Self-employed	*108*	*2.84%*
Family business worker	*685*	*18.03%*
Fishing	61	1.61%
Fishing, employed	*17*	*0.45%*
Fishing, self-employed	*44*	*1.16%*
Employee	89	2.34%
Casual labour (ganyu)	24	0.63%
Student	4	0.11%
Unemployed, not seeking work	9	0.24%
Other	14	0.37%

Notes: Occupation as measured during the 2013 member census, N=3,801 members. 519 members were no longer active, but are included to avoid selection bias. Analysis is conducted with and without these individuals, see Section 7. Missing observations reflect answers of "do not know" or "not applicable". Occupation denotes an individual's primary economic activity, if engaged in multiple activities.

Table 14: Test of representativeness – matched dyads vs. 2013 census dyads

	Mean All Dyads	Mean Matched Dyads	Diff.	Std. Error	T stat	Full N	Matched N
Group membership							
Same VSLA group	0.17	0.30	-0.13***	(0.01)	-24.27	289914	7309
Occupation							
Same economic activity	0.56	0.71	-0.15***	(0.01)	-27.11	289914	7309
Absolute differences							
Male	0.38	0.39	-0.02**	(0.01)	-2.89	289467	7309
Female-headed household	0.32	0.34	-0.03***	(0.01)	-4.64	288740	7309
Age	12.64	12.17	0.46***	(0.12)	3.94	286763	7309
Some post-primary educ.	0.33	0.26	0.07***	(0.01)	13.29	289914	7309
Literate (read & understand newspaper)	0.26	0.29	-0.03***	(0.01)	-5.12	288342	7309
Father well-off in village (scale 1-5)	1.41	1.42	-0.00	(0.01)	-0.12	253041	6415
Spouse's father well-off in village (scale 1-5)	1.43	1.40	0.03*	(0.01)	2.44	252485	6954
Household well-off in group (scale 1-9)	1.40	1.29	0.11***	(0.01)	8.31	284485	7309
HH owns a bicycle	0.46	0.46	0.01	(0.01)	1.21	288740	7309
# Goats	1.84	2.30	-0.45***	(0.04)	-11.14	288240	7309
Sums							
Farmer	1.37	1.61	-0.24***	(0.01)	-33.75	289914	7309
Businessperson	0.48	0.32	0.16***	(0.01)	24.67	289914	7309
Male	0.53	0.55	-0.02**	(0.01)	-2.68	289467	7309
Female-headed household	0.40	0.45	-0.05***	(0.01)	-6.69	288740	7309
Age	72.12	77.04	-4.92***	(0.19)	-25.52	286763	7309
Some post-primary educ.	0.43	0.31	0.11***	(0.01)	18.76	289914	7309
Literate (read & understand newspaper)	1.68	1.63	0.04***	(0.01)	6.87	288342	7309
Father well-off in village (scale 1-5)	6.76	6.64	0.12***	(0.02)	5.11	253041	6415
Spouse's father well-off in village (scale 1-5)	6.85	6.87	-0.02	(0.02)	-0.84	252485	6954
Household well-off in group (scale 1-9)	15.04	15.41	-0.37***	(0.02)	-18.06	284485	7309
HH owns a bicycle	0.99	1.09	-0.10***	(0.01)	-12.09	288740	7309
# Goats	2.44	3.32	-0.89***	(0.05)	-16.90	288240	7309

Notes: All variables from the 2013 member census, N=3,801 members. N=722 of these members could be matched by name to the 2009-11 panel dataset. All possible dyads in which both individuals live in the same village are constructed for the full sample (N=289,914) and for the matched subsample (N=7,301). Missing observations reflect answers of "do not know" or "not applicable". Occupation denotes an individual's primary economic activity, if engaged in multiple activities.

Table 15: Dyadic regressions – 2013 member census and matched subsample

	(1) **Full CSAE sample** Mfx / (s.e.)	(2) **Matched subsample** Mfx / (s.e.)	P-value $\beta_1 = \beta_2$
Occupation			
Same economic activity	0.086***	0.206***	0.132
	(0.008)	(0.045)	
Absolute differences			
Male	-0.053***	-0.034	0.054*
	(0.015)	(0.028)	
Female-headed household	-0.012***	-0.065***	0.034**
	(0.004)	(0.023)	
Age	-0.001***	-0.001	0.629
	(0.000)	(0.001)	
Some post-primary educ.	-0.008	-0.005	0.724
	(0.005)	(0.021)	
Literate (read & understand newspaper)	-0.013*	-0.009	0.417
	(0.0067)	(0.018)	
Father well-off in village (scale 1-5)	-0.013***	-0.025***	0.577
	(0.002)	(0.010)	
Spouse's father well-off in village (scale 1-5)	-0.013***	-0.030**	0.379
	(0.003)	(0.012)	
Household well-off in group (scale 1-9)	-0.010***	-0.020**	0.659
	(0.003)	(0.010)	
HH owns a bicycle	-0.003	0.009	0.159
	(0.003)	(0.009)	
# Goats	-0.006***	-0.012***	0.407
	(0.002)	(0.004)	
Sums	✓	✓	
Village f.e.'s	✓	✓	
Observations	219747	5878	
Pseudo R^2	0.129	0.142	
Baseline predicted probability	0.161	0.168	

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. All variables from the 2013 member census, N=3,801 members. N=722 of these members could be matched by name to the 2009-11 panel dataset. All possible dyads in which both individuals live in the same village are constructed for the full sample (N=289,914) and for the matched subsample (N=7,301). Missing observations reflect answers of "do not know" or "not applicable". Occupation denotes an individual's primary economic activity, if engaged in multiple activities. Reported effects are marginal effects estimated at the mean.

Table 16: Dyadic regressions – treated vs. control villages, and villages with above vs. below median number of groups

	(1) **Treated 2009** Mfx / (s.e.)	(2) **Treated 2011** Mfx / (s.e.)	(3) **> 2 groups** Mfx / (s.e.)	(4) **≤ 2 groups** Mfx / (s.e.)
Occupation				
Same economic activity	0.074***	0.139***	0.079***	0.144***
	(0.007)	(0.028)	(0.007)	(0.040)
Absolute differences				
Male	-0.056***	-0.032**	-0.052***	-0.035
	(0.014)	(0.013)	(0.014)	(0.026)
Female-headed household	-0.008*	-0.023**	-0.007*	-0.068**
	(0.005)	(0.011)	(0.004)	(0.029)
Age	-0.001***	-0.001***	-0.001***	-0.002**
	(0.000)	(0.000)	(0.000)	(0.001)
Some post-primary educ.	-0.008	-0.006	-0.007	-0.013
	(0.005)	(0.011)	(0.005)	(0.021)
Literate (read & understand newspaper)	-0.012	-0.019	-0.010	-0.049*
	(0.008)	(0.015)	(0.007)	(0.027)
Father well-off in village (scale 1-5)	-0.010***	-0.022***	-0.011***	-0.030***
	(0.002)	(0.003)	(0.001)	(0.010)
Spouse's father well-off in village (scale 1-5)	-0.010***	-0.022***	-0.011***	-0.032***
	(0.003)	(0.008)	(0.003)	(0.012)
Household well-off in group (scale 1-9)	-0.008***	-0.017***	-0.009***	-0.010
	(0.003)	(0.005)	(0.003)	(0.009)
HH owns a bicycle	-0.003	-0.001	-0.004	0.010
	(0.003)	(0.010)	(0.003)	(0.007)
# Goats	-0.004**	-0.011***	-0.005***	-0.011**
	(0.002)	(0.003)	(0.002)	(0.004)
Sums	✓	✓	✓	✓
Village f.e.'s	✓	✓	✓	✓
Observations	171738	48009	204089	15658
Pseudo R^2	0.122	0.085	0.086	0.019
Baseline predicted probability	0.217	0.206	0.141	0.494

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. All variables from the 2013 member census, N=3,801 members. 519 of these members were no longer active, but are included here to avoid selection bias. Results are robust to excluding these individuals, see Section 7. All possible dyads in which both individuals live in the same village are constructed, N=289,914. Missing observations reflect answers of "do not know" or "not applicable". Occupation denotes an individual's primary economic activity, if engaged in multiple activities. Reported effects are marginal effects estimated at the mean.

Table 17: Dyadic regressions – matched subsample, treated vs. control villages, and villages with above vs. below median number of groups

	(1) **Treated 2009** Mfx / (s.e.)	(2) **Treated 2011** Mfx / (s.e.)	(3) **> 2 groups** Mfx / (s.e.)	(4) **≤ 2 groups** Mfx / (s.e.)
Occupation				
Same economic activity	0.178**	0.212	0.106	0.295
	(0.087)	(0.160)	(0.067)	(0.191)
Absolute differences				
Present-biased	1.173***	0.073	0.899***	0.153
	(0.051)	(0.076)	(0.036)	(0.103)
Male	0.020	-0.014	0.001	-0.047
	(0.034)	(0.039)	(0.027)	(0.107)
Female-headed household	-0.274***	0.110	-0.136**	-0.137
	(0.061)	(0.077)	(0.054)	(0.171)
Age	-0.003	-0.014**	-0.002	-0.015***
	(0.003)	(0.006)	(0.002)	(0.006)
Some post-primary educ.	-0.045	-0.040	-0.005	-0.135*
	(0.081)	(0.069)	(0.056)	(0.072)
Literate (read & understand newspaper)	-0.028	-0.002	-0.015	-0.107
	(0.039)	(0.154)	(0.022)	(0.227)
Father well-off in village (scale 1-5)	-0.014	-0.030	-0.009	-0.048
	(0.012)	(0.035)	(0.009)	(0.034)
Spouse's father well-off in village (scale 1-5)	-0.050***	-0.007	-0.032***	-0.030
	(0.015)	(0.017)	(0.012)	(0.033)
Household well-off in group (scale 1-9)	-0.060***	0.029	-0.023	-0.085*
	(0.013)	(0.033)	(0.015)	(0.049)
HH owns a bicycle	-0.005	0.044	0.003	-0.077
	(0.034)	(0.056)	(0.022)	(0.080)
# Goats	-0.006	-0.043*	-0.010**	-0.014
	(0.006)	(0.026)	(0.005)	(0.033)
Sums	✓	✓	✓	✓
Village f.e.'s	✓	✓	✓	✓
Observations	920	372	944	348
Pseudo R^2	0.271	0.227	0.147	0.194
Baseline predicted probability	0.331	0.298	0.181	0.554

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. N=722 individuals are matched from the 2013 census to the 2009-11 panel data. N=383 of these individuals are matched to the longer panel survey including preference modules. All possible dyads in which both individuals live in the same village are constructed, N=7,314 for the general survey and N=1,641 for the full survey including preference modules. Missing values reflect "do not know", "not applicable", or inconsistent answers in the case of risk preferences. Time preferences are taken from the 2010 survey, wave since the 2009 wave did not include the far frame for females. Present-biased (future-biased) is a dummy equal to one if the response to the near frame is more impatient (patient) than the response to the far frame. Reported effects are marginal effects estimated at the mean.

Table 18: Dyadic regressions – 2013 member census, weighted

	(1) **Matched subsample** Mfx / (s.e.)
Occupation	
Same economic activity	0.098***
	(0.035)
Absolute differences	
Present-biased	0.096***
	(0.032)
Male	0.000
	(0.014)
Female-headed household	-0.094***
	(0.032)
Age	-0.004***
	(0.001)
Some post-primary educ.	-0.013
	(0.025)
Literate (read & understand newspaper)	-0.035
	(0.028)
Father well-off in village (scale 1-5)	-0.009
	(0.006)
Spouse's father well-off in village (scale 1-5)	-0.016**
	(0.007)
Household well-off in group (scale 1-9)	-0.027***
	(0.008)
HH owns a bicycle	0.008
	(0.016)
# Goats	-0.005
	(0.004)
Sums	✓
Village f.e.'s	✓
Observations	1280
Pseudo R^2	0.280
Baseline predicted probability	0.197

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. N=722 individuals are matched from the 2013 census to the 2009-11 panel data. N=383 of these individuals are matched to the longer panel survey including preference modules. All possible dyads in which both individuals live in the same village are constructed, N=7,314 for the general survey and N=1,641 for the full survey including preference modules. Missing values reflect "do not know", "not applicable", or inconsistent answers in the case of risk preferences. Time preferences are taken from the 2010 survey, wave since the 2009 wave did not include the far frame for females. "Present-biased" ("future-biased") is a dummy equal to one if the response to the near frame is more impatient (patient) than the response to the far frame. Reported effects are marginal effects estimated at the mean.

Table 19: Dyadic regressions – time preference measures, matched subsample, weighted

	(1) **Pr(same group= 1)** Mfx / (s.e.)	(2) **Pr(same group= 1)** Mfx / (s.e.)	(3) **Pr(same group= 1)** Mfx / (s.e.)	(4) **Pr(same group= 1)** Mfx / (s.e.)	(5) **Pr(same group= 1)** Mfx / (s.e.)
Occupation					
Same economic activity	0.098***	0.093**	0.098***	0.102***	0.097***
	(0.035)	(0.038)	(0.034)	(0.034)	(0.035)
Absolute differences					
Present-biased	0.096***		0.099***	0.097***	0.096***
	(0.032)		(0.031)	(0.030)	(0.032)
Future-biased					0.002
					(0.015)
Minimum switch-point, near frame		0.000	0.001		
		(0.002)	(0.002)		
Minimum switch-point, far frame		0.000	-0.001		
		(0.003)	(0.003)		
Patience above median, near frame				0.015	
				(0.017)	
Patience above median, far frame				-0.022	
				(0.016)	
Wealth & income controls (abs. diffs)	✓	✓	✓	✓	✓
Demographic controls (abs. diffs)	✓	✓	✓	✓	✓
Sums	✓	✓	✓	✓	✓
Village f.e.'s	✓	✓	✓	✓	✓
Observations	1280	1280	1280	1280	1280
Pseudo R^2	0.280	0.272	0.282	0.286	0.281
Baseline predicted probability	0.197	0.196	0.197	0.197	0.197

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. N=722 individuals are matched from the 2013 census to the 2009-11 panel data. N=383 of these individuals are matched to the longer panel survey including preference modules. All possible dyads in which both individuals live in the same village are constructed, N=7,314 for the general survey and N=1,641 for the full survey including preference modules. Missing values reflect "do not know", "not applicable", or inconsistent answers in the case of risk preferences. All variables presented here were measured in the 2009 wave of the panel survey, except time preferences, which are taken from the 2010 survey wave since the 2009 wave did not include the far frame for females. "Present-biased" ("future-biased") is a dummy equal to one if the response to the near frame is more impatient (patient) than the response to the far frame. Minimum switch-point is the lower bound of the interval in which the respondent switched to preferring the payment dated one month later compared to a 2000 MK payment on the earlier date. 150 MK $\approx$ 1 USD at the time of the 2009 and 2010 surveys. Patience above median is a dummy for having a switch-point below the median in that frame. Reported effects are marginal effects estimated at the mean.

Table 20: Dyadic regressions – effects of controls, matched subsample, weighted

	(1) **Pr(same group= 1)** Mfx / (s.e.)	(2) **Pr(same group= 1)** Mfx / (s.e.)	(3) **Pr(same group= 1)** Mfx / (s.e.)	(4) **Pr(same group= 1)** Mfx / (s.e.)	(5) **Pr(same group= 1)** Mfx / (s.e.)	(6) **Pr(same group= 1)** Mfx / (s.e.)	(7) **Pr(same group= 1)** Mfx / (s.e.)	(8) **Pr(same group= 1)** Mfx / (s.e.)	(9) **Pr(same group= 1)** Mfx / (s.e.)	(10) **Pr(same group= 1)** Mfx / (s.e.)
Occupation										
Same economic activity	0.098***		0.117**	0.119**	0.088**	0.094***	0.098***	0.158*	0.105**	0.113**
	(0.035)		(0.049)	(0.050)	(0.038)	(0.035)	(0.035)	(0.084)	(0.041)	(0.055)
Absolute differences										
Present-biased	0.096***	0.093***	0.052**	0.057**		0.117***	0.096***	0.187**	0.065**	0.057*
	(0.032)	(0.028)	(0.022)	(0.023)		(0.035)	(0.031)	(0.075)	(0.028)	(0.033)
Risk-neutral			0.026							
			(0.036)							
Risk aversion above median				-0.020						-0.017
				(0.015)						(0.017)
Ever hides money from spouse					-0.002	-0.005				0.010
					(0.019)	(0.015)				(0.018)
Female HH decision-making power (index 0-8)					0.005	0.003				0.016**
					(0.007)	(0.007)				(0.008)
HH important in village decisions (scale 1-6)							-0.013			-0.017
							(0.009)			(0.012)
HH speaks at village meetings							-0.001			0.035**
							(0.012)			(0.014)
GPS distance between hhs (km)								-0.072**		-0.048
								(0.033)		(0.033)
Monthly consumption per capita, 1000 MK									-0.009	-0.008
									(0.014)	(0.016)
Wealth & income controls (abs. diffs)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls (abs. diffs)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sums	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Village f.e.'s	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1280	1280	1021	1021	969	947	1265	1249	1147	649
Pseudo R^2	0.280	0.264	0.270	0.271	0.242	0.268	0.285	0.252	0.267	0.343
Baseline predicted probability	0.197	0.191	0.189	0.189	0.193	0.197	0.194	0.294	0.177	0.200

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. N=722 individuals are matched from the 2013 census to the 2009-11 panel data. N=383 of these individuals are matched to the longer panel survey including preference modules. All possible dyads in which both individuals live in the same village are constructed, N=7,314 for the general survey and N=1,641 for the full survey including preference modules. Missing values reflect "do not know", "not applicable", or inconsistent answers in the case of risk preferences. All variables presented here were measured in the 2009 wave of the panel survey, except time preferences, which are taken from the 2010 survey wave since the 2009 wave did not include the far frame for females. "Present-biased" ("future-biased") is a dummy equal to one if the response to the near frame is more impatient (patient) than the response to the far frame. Risk-neutral is a dummy equal to one if the respondent prefers a 50-50 lottery to its expected value for certain, and thus could indicate risk-neutral or risk-seeking behaviour. Female HH decision-making power is constructed from questions over four types of economic decisions, scoring one if the female has some say in the decision and two if she has complete control. GPS co-ordinates are recorded by the enumerator when visiting the household. Monthly consumption per capita is estimated from a full consumption roster and local prices using USAID's Poverty Assessment Tool for Malawi. Reported effects are marginal effects estimated at the mean.

Table 21: Individual financial outcomes – full member census, treated vs. control villages

	(1) **Total loans (MK), treated 2009** β / (s.e.)	(2) **Total loans (MK), treated 2011** β / (s.e.)	(3) **Total shares, treated 2009** β / (s.e.)	(4) **Total shares, treated 2011** β / (s.e.)
Occupation				
Farmer	-15018.40***	-11376.07**	3.95	19.65**
	(4900.61)	(4576.70)	(4.84)	(9.34)
% Farmer in group	-42206.81	-58949.44	37.61	23.67
	(41325.04)	(83020.03)	(32.94)	(154.80)
Village f.e.'s	✓	✓	✓	✓
Observations	2181	1201	2181	1201
Mean dep. var.	56706.92	59228.13	189.13	225.52
R^2	0.025	0.014	0.004	0.005

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. All dependent variables are constructed from the individual passbook data. We recover passbooks for N=3,382 members. Missing observations reflect answers of "do not know" or "not applicable". Total borrowed, total interest paid, total shares bought are summed for each member across all weeks of the current cycle. Occupation and occupation of other groups members are measured during the full 2013 member census, N=3,801 members. 519 members were no longer active, but are included to avoid selection bias. Analysis is conducted with and without these individuals, see Section 7. Occupation denotes an individual's primary economic activity, if engaged in multiple activities. Treated villages received the VSLA program in 2009, control villages in 2011.

Table 22: Individual financial outcomes – matched subsample

	(1) **Total loans (MK)** β / (s.e.)	(2) **Total interest(MK)** β / (s.e.)	(3) **Total shares** β / (s.e.)	(4) **Total savings(MK)** β / (s.e.)	(5) **Loan:savings ratio** β / (s.e.)	(6) **Net financial gain(MK)** β / (s.e.)
Occupation						
Farmer	-12738.41	-436.92	-8.18	-2738.83	-2.48	-1609.88
	(11763.23)	(2650.33)	(28.12)	(4461.70)	(2.41)	(2401.78)
Present-biased	-21525.01***	1820.78	16.15	2455.43	-0.89**	247.33
	(6382.63)	(2968.52)	(23.56)	(4020.99)	(0.38)	(4015.22)
Poverty						
Monthly consumption per capita, 1000 MK	-10657.85**	-1882.03	-2.96	29.98	-0.52**	1536.00
	(4885.47)	(1667.47)	(11.04)	(1415.58)	(0.25)	(1269.38)
Food security poor (dummy)	-9982.66	-2763.65	-20.33	-4561.66**	-0.70	-526.73
	(8566.00)	(1935.96)	(19.20)	(2226.24)	(1.38)	(1750.46)
Village f.e.'s	✓	✓	✓	✓	✓	✓
Observations	297	297	297	297	288	292
Mean dep. var.	57024.49	11370.96	209.65	25172.83	2.84	-753.43
R^2	0.029	0.006	0.007	0.017	0.027	0.005

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. All dependent variables are constructed from the individual passbook data. We recover passbooks for N=3,382 members, of whom 722 can be matched to the panel data containing measures of present-bias, income and food security. Missing observations reflect answers of "do not know" or "not applicable". Total loans borrowed, total interest paid, and total shares bought are summed for each member across all weeks of the current cycle. Total savings is calculated by multiplying total shares by the share price. Loan:savings ratio is calculated by dividing total borrowed by total savings. Net gain is calculated as the difference between predicted payout on purchased shares and interest charged on loans received. Occupation and occupation of other groups members are measured during the full 2013 member census, N=3,801 members. Occupation denotes an individual's primary economic activity, if engaged in multiple activities. Present-biased [time-consistent] (future-biased) is a dummy equal to one if the response to the near frame is more impatient than [the same as] (more impatient than) than the response to the far frame in the panel survey time preference activity. Monthly consumption per capita is estimated from a full consumption roster and local prices using USAID's Poverty Assessment Tool for Malawi. Malawi's GNI per capita in 2009 was $26.6/month, but these are particularly poor households in a very remote region. Food security poor is equal to one if the household reports consuming fewer than three meals yesterday.

Table 23: Individual financial outcomes – matched subsample, present-biased or time-consistent individuals only

	(1) **Total loans (MK)** β / (s.e.)	(2) **Total interest(MK)** β / (s.e.)	(3) **Total shares** β / (s.e.)	(4) **Total savings(MK)** β / (s.e.)	(5) **Loan:savings ratio** β / (s.e.)	(6) **Net financial gain(MK)** β / (s.e.)
Occupation						
Farmer	-22989.18	-1793.95	-14.48	-4621.47	-3.59	-42.54
	(16602.30)	(4342.63)	(39.36)	(6249.91)	(3.64)	(3730.40)
Present-biased	-30802.32***	-131.00	14.00	1083.41	-1.26**	1965.95
	(11899.12)	(3432.97)	(26.77)	(4306.75)	(0.61)	(4241.19)
Poverty						
Monthly consumption per capita, 1000 MK	-10902.69*	-1842.40	0.68	566.28	-0.46*	1949.57
	(6299.26)	(1958.55)	(12.71)	(1761.44)	(0.24)	(1575.02)
Food security poor (dummy)	-19969.67	-5761.49*	-28.62	-6349.36**	-1.16	1637.32
	(12683.81)	(3004.65)	(20.18)	(2742.04)	(2.14)	(2738.63)
Village f.e.'s	✓	✓	✓	✓	✓	✓
Observations	217	217	217	217	210	213
Mean dep. var.	58136.29	11988.67	214.46	25857.83	3.03	-1470.23
R^2	0.048	0.010	0.013	0.032	0.038	0.005

Notes: *, ** and *** represent p<0.10, p<0.05 and p<0.01 respectively. All dependent variables are constructed from the individual passbook data. We recover passbooks for N=3,382 members, of whom 722 can be matched to the panel data containing measures of present-bias, income and food security. Missing observations reflect answers of "do not know" or "not applicable". Total loans borrowed, total interest paid, and total shares bought are summed for each member across all weeks of the current cycle. Total savings is calculated by multiplying total shares by the share price. Loan:savings ratio is calculated by dividing total borrowed by total savings. Net gain is calculated as the difference between predicted payout on purchased shares and interest charged on loans received. Occupation and occupation of other groups members are measured during the full 2013 member census, N=3,801 members. Occupation denotes an individual's primary economic activity, if engaged in multiple activities. Present-biased [time-consistent] (future-biased) is a dummy equal to one if the response to the near frame is more impatient than [the same as] (more impatient than) than the response to the far frame in the panel survey time preference activity. Monthly consumption per capita is estimated from a full consumption roster and local prices using USAID's Poverty Assessment Tool for Malawi. Malawi's GNI per capita in 2009 was $26.6/month, but these are particularly poor households in a very remote region. Food security poor is equal to one if the household reports consuming fewer than three meals yesterday.

A.3 Testing for assortative matching using dyadic regressions

In this Appendix, we demonstrate how testing for negative or positive assortative matching in groups can be implemented using a dyadic regression. We start with a simple numerical example to illustrate the intuition behind the proposed method. We then offer various simulations to show how the method performs in various situations.

A.3.1 A simple example

Imagine a population with four individuals numbered 1 to 4. Individuals 1 and 2 are high types; 3 and 4 are low types. The difference between high and low type is 1. There are six dyads in this dataset:

Table A1. Possible dyads with difference in type

Individual 1	Individual 2	Difference in type
1	2	0
1	3	1
1	4	1
2	3	1
2	4	1
3	4	0

These four individuals can form two groups of two individuals. Since order is irrelevant, there are three possible groupings:

Table A2. Possible groupings with average differece within and across groups

Pair 1	Pair 2	Av. Diff. within groups	Av. Difference across groups
$1-2$	$3-4$	0	1
$1-3$	$2-4$	1	0.5
$1-4$	$2-3$	1	0.5

In each grouping, two of the six dyads are within groups, and four are across groups. In the first grouping, individuals are sorted by type and the difference in type for grouped dyads $1-2$ and $3-4$ is 0. As shown in Table A1, the difference across each the four ungrouped dyads is 1. In the other two groupings, the difference for within group dyads is 1 while the average difference for across group dyads is 0.5.

Now imagine that we have a data composed of N pools of four individuals grouped in pairs. Within each pool we form all possible six dyads, giving us $6N$ observations in total: $2N$ within groups and $4N$ across groups. Let d_{ijk} be the difference in type between individuals i and j in pool k. Further let $g_{ijk} = 1$ if i and j are in a group and 0 otherwise. The estimated regression is:

$$d_{ijk} = \alpha + \beta g_{ijk} + u_{ijk}$$

It immediately follows that if all groups in all pools are $1-2$ and $3-4$, regressing d_{ijk} on g_{ijk} will yield a coefficient of -1, which is the average difference within groups minus the average difference across groups. Similarly if all groups are of the form $1-3$ and $2-4$ or $1-4$ and

$2-3$, then the coefficient of $g_{\iota jk}$ will be $1-0.5=0.5$. In other words, perfect positive assorting – the first case – yields $\widehat{\beta}<0$ while negative assorting – the second case – yields $\widehat{\beta}>0$.

Let us now generalize this example to imperfect, i.e., probabilistic assorting. We consider three possible assorting rules: neutral; positive assorting; and negative assorting. Examples of assorting probabilities are shown in Table A3, each with its associated $E[d_{ijk}]$:

Table A3. Expected dyadic regression coefficient for different assorting probabilities

Pair 1	Pair 2	Neutral	$E[d_{ijk}]$	Positive	$E[d_{ijk}]$	Negative	$E[d_{ijk}]$
$1-2$	$3-4$	0.333	-0.333	0.533	-0.533	0.133	-0.133
$1-3$	$2-4$	0.333	0.167	0.233	0.117	0.433	0.217
$1-4$	$2-3$	0.333	0.167	0.233	0.117	0.433	0.217
		$E[\widehat{\beta}]=$	0	$E[\widehat{\beta}]=$	-0.3	$E[\widehat{\beta}]=$	0.3

We see that, as it should be, neutral assorting yields $E[\widehat{\beta}]=0$ while imperfect positive and negative assorting yields $E[\widehat{\beta}]$ values that are negative and positive, respectively (e.g., Fafchamps and Gubert, 2007).

The above example can be turned on its head to ask whether differences in type *predict* groupings. This is achieved by estimating a predictive regression of the form:

$$g_{ijk}=\theta+\gamma d_{ijk}+e_{ijk} \tag{2}$$

Since both β and γ capture the correlation between g_{ijk} and δ_{ijk}, it follows that the sign properties of $\widehat{\beta}$ transfer to $\widehat{\gamma}$: if similarity ot type predicts a dyad being in the same group, large differences in type will predict *not* being in a group – hence a negative $\widehat{\gamma}$.

With this transformation, the example can easily be expanded to a situation in which there are two or more dyadic variables susceptible of predict groupings. For example, imagine that individuals can also be divided into male and female. Suppose for instance that individuals match negatively on gender but positively on type, then knowing that a dyad has one male and one female of the same type predicts a higher grouping probability $E[g_{ijk}]$. If gender and type are correlated, estimating equation (2) will help disentangle the respective predictive role of gender and type. For it will let us test whether observed differences in types across pairs are a by-product from matching on gender rather than a manifestation of assorting on type.

A.3.2 Simulations

The above apparatus can be extended to groups of arbitrary size. We illustrate this with a couple of simulations below. The details of the implementation are given in the next subsection. Neutral assorting means sorting on an unobserved variable only. Positive assortative matching (PAM) is here defined as matching on both gender and type, in addition to the unobserved variable; hybrid assortative matching (HAM) is defined as negative assorting on gender but positive assorting on type and on the unobserved variable. Gender and type are uncorrelated by design. Table A4 considers small groups of 3 individuals in pools of 15; Table A4 consideres larger groups of 10 individuals in pools of 20. Both Tables illustrates how dyadic estimation correctly identifies the correct sign of the gender and type dummies – i.e., positive in case of negative assorting and vice versa.

Table A4. Dyadic estimation of simulated data – group of size 3

	Neutral	S.E.	PAM	S.E.	HAM	S.E.
Gender dif.	−0.015	0.016	−0.179***	0.010	0.074***	0.003
Type dif.	−0.003	0.010	−0.048***	0.014	−0.079***	0.015
Intercept	0.151	0.093	0.258	0.006	0.137	0.007
N.obs.	6300		6300		6300	

Table A4 notes: Size of group: 3. Number of groups per pool: 5. Number of pools: 30. PAM is positive assorting on gender and type. NAM is negative assorting on gender and positive assorting on type. Estimator is OLS with standard errors clustered by pool. Results shown are for one simulation.

Table A5. Dyadic estimation of simulated data – group of size 10

	Neutral	S.E.	PAM	S.E.	HAM	S.E.
Gender dif.	0.011	0.013	−0.285***	0.022	0.045***	0.003
Type dif.	−0.003	0.011	−0.056***	0.017	−0.186***	0.014
Intercept	0.469	0.007	0.644	0.011	0.523	0.006
N.obs.	11400		11400		10400	

Table A5 notes: Size of group: 10. Number of groups per pool: 2. Number of pools: 30. PAM is positive assorting on gender and type. NAM is negative assorting on gender and positive assorting on type. Estimator is OLS with standard errors clustered by pool. Results shown are for one simulation.

A.4 Implementation details

The simulations are implemented as follows. For each pair, we create a value of the match variable that takes three possible values: neutral V_{ij}^n; PAM V_{ij}^p; and HAM V_{ij}^h. They are calculated as follows:

$$\begin{aligned} V_{ij}^n &= |e_i - e_j| \\ V_{ij}^p &= -|gender_i - gender_j| - 0.5|type_i - type_j| + |e_i - e_j| \\ V_{ij}^h &= |gender_i - gender_j| - 0.5|type_i - type_j| + |e_i - e_j| \end{aligned}$$

where e_i and e_j are independent random draws from a standard normal distribution, $gender = 1$ for 55% of randomly selected individuals and $type = 1$ for 75% of randomly selected individuals. Type and gender are independently assigned.

Values are then used to calculate the aggregate value of any grouping. To find the groupings yielding the highest average value on the match, we randomly try 100 different groupings and keep the grouping with the highest aggregate value across all groups. This is done separately for each pool. In practice, most of the improvement in value is achieved after a small number of iterations, e.g., less than 10. Small incremental improvements are occasionally observed up to 50 iterations, rarely above that. We are therefore confident that the simulated groupings are arbitrarily close to 'equilibrium' groupings that would maximize aggregate pay-offs within each pool.