Online Appendix for Inequality and Growth in China

Haiyan Lin and Markus Brueckner

More Regression Results and Discussions

1. OLS Estimates

Appendix Table 1 reports OLS estimates. The inequality variables included on the right-hand side of the estimating equation are the same as in Table 10. Panel A of Appendix Table 1 reports pooled OLS estimates, i.e. without county fixed effects as controls; Panel B of Appendix Table 1 reports OLS estimates of a model that includes county fixed effects. For all estimates reported in Panels A and B, the model includes time fixed effects.

Comparing Appendix Table 1 and Table 10, one can see that OLS estimates are qualitatively the same as sys-GMM estimates. Quantitatively, OLS yields estimated coefficients on inequality and the interaction term that are, in absolute value, smaller than the corresponding sys-GMM estimates. One reason for this is classical measurement error that attenuates OLS estimates towards zero. Measurement error that is serially uncorrelated does not attenuate sys-GMM estimates. Another reason for the larger sys-GMM estimates is reverse causality: according to the Galor and Zeira (1993) model an increase in income per capita reduces inequality. OLS estimates of the effect that inequality has on growth are thus downward biased due to reverse causality going from higher income to less inequality.

2. GMM Estimates of Models Without an Interaction Between Inequality and Initial Income

The conventional approach in the empirical literature, starting with Forbes (2000), and see Kraay (2015) for a review of recent empirical papers, was to estimate models of the relationship between inequality and growth without interaction terms. That is, in the majority of empirical papers the endeavor was to estimate so-called average effects of inequality on growth. To see what we obtain for our sample of Chinese counties if we follow the conventional approach, we report in Appendix Table 2 estimates of a model that does not include an interaction between the Gini and counties' initial household incomes per capita. We note upfront that Chinese counties' incomes per capita have grown substantially during 1989-2015. Given the sys-GMM estimates in Table 2 that showed a significant interaction between initial income and the Gini, the purpose of reporting results here from a model without an interaction term is simply to see what this yields for the average effect of inequality on growth. Clearly, given that Chinese counties' incomes per capita are currently much larger than what they were on average during 1989-2015, this average effect is not informative with regard to the question of what is the effect of inequality on economic growth for current levels of development. Nevertheless, it is worthwhile to report and discuss these estimates here in order to see what one obtains by following the conventional approach.

One can see from Appendix Table 2 that, across all columns, the estimated coefficient on the Gini is negative, and not significantly different from zero at the conventional significance levels. The estimated average effect of the Gini on economic growth in Appendix Table 2, is quantitatively similar to what the estimates of the interaction model in Table 2 show with regard to the effect that the Gini has on economic growth for initial income per capita equal to the sample mean. In the estimates of the interaction model, the effect of inequality on growth is $(\beta+\gamma*\ln y_{i,t-1})$. To obtain in the interaction model the average effect of inequality on growth, one has to set initial household income per capita equal to the sample mean: i.e. in the interaction model the average effect of inequality on growth is $(\beta+\gamma*\ln y_{mean})$.

According to the estimates in column (1) of Table 2, the effect of a 1 percentage point increase in the Gini on annual growth, evaluated at sample mean initial household income per capita, is -0.07 percentage points; the standard error associated with this effect is 0.08. For the winsorized and limited lag estimates in columns (2) and (3) of Table 2 the effects, again evaluated at mean initial household income per capita, are -0.06 and -0.04, respectively, with standard errors of

0.09 and 0.09, respectively. For the one-step sys-GMM estimates in columns (4)-(6) of Table 2 the effects at sample mean household income per capita are -0.04, -0.05, and -0.08, respectively, with standard errors 0.08, 0.09, and 0.09.

Looking at Appendix Table 2, one can see that the average effects of the Gini on growth, that are obtained from a model without interaction term, range between -0.02 to -0.08. These estimated average effects of inequality on growth are indeed very similar to the effects of inequality on growth, discussed in the paragraph above, from the interaction model when initial income per capita is set equal to the sample mean; those effects range from -0.04 to -0.08.

3. **A Model with a Quadratic Inequality Term**

In Appendix Table 3 we report sys-GMM estimates of a model that includes a quadratic inequality term. In the empirical literature, a quadratic inequality term was included on the right-hand side of the econometric model to investigate whether there is an inverted U-shaped relationship between inequality and growth.

The estimates in Appendix Table 3, which are based on strong and valid instruments, show that the estimated coefficient on the quadratic inequality term is, in the majority of specifications, not significantly different from zero. The exception is the 75-25 ratio: i.e. in column (4) of Appendix Table 3, one can see that the coefficient on the squared term of the 75-25 income ratio is significantly negative while the coefficient on the linear term of the 75-25 income ratio is significantly positive.

Qualitatively, the signs of the estimated coefficients are, for the majority of specifications, consistent with an inverted U-shaped relationship between inequality and growth. For example, in column (1) of Appendix Table 3, the estimated coefficient (standard error) on $Gini_{it-1}$ is 2.42 (1.17); the coefficient (standard error) on $Gini_{it-1}^2$ is -0.16 (0.27); and the coefficient (standard error) on $Gini_{it-1}*lny_{it-1}$ is -0.27 (0.13). One can reject the joint hypothesis that the estimated coefficients on

$Gini_{it-1}$, $Gini_{it-1}^2$, and $Gini_{it-1}*lny_{it-1}$ are equal to zero at the 10 percent level (p-value 0.09).

For the estimates in column (1) of Appendix Table 3 the marginal effect of the Gini on growth is: $2.42 – 0.32*Gini_{it-1} – 0.27*lny_{it-1.}$ We now interpret this marginal effect. Consider two points in time: (1) at the beginning of the sample (1989); and (2) at the end of the sample (2015). For an initial income equal to the sample mean in the year 1989, the marginal effect of the Gini on growth is: $0.26 – 0.32*Gini$. The marginal effect of the Gini on economic growth is positive up to the point where the Gini is 0.82; above 0.82 the marginal effect is negative. In the year 1989, all counties in the sample had a Gini of less than 0.82. Thus, for all sample values of the Gini in the year 1989 an increase in inequality had a positive effect on economic growth. Increases in inequality had larger positive effects on economic growth in counties with initially more equal distributions of income.

For an initial income equal to the sample mean in the year 2015 the marginal effect of the Gini on growth is: -0.28 – 0.32*Gini. The marginal effect of an increase in the Gini on economic growth is negative for all values of the Gini. Increases in the Gini had larger negative effects on economic growth in counties with initially more unequal distributions of income.

4. The Benjamin et al. (2011) Dataset

In this section we discuss sys-GMM estimates using the panel data of Benjamin et al. (2011). The Benjamin et al. (2011) panel data covers 82 villages in 9 provinces in China during 1987 to 2002.

One important issue in the comparison of estimates between the Benjamin et al. dataset and our CHNS dataset is the time period. The Benjamin et al. panel dataset covers the period 1987-2002. The estimates that we had discussed in previous sections for the CHNS dataset were for the longest possible time period given the available CHNS survey data, which is 1989-2015. As discussed in Section 2, household income per capita of Chinese counties has grown substantially during 1989-2015. For the purpose of comparing estimates of the econometric models for the

Benjamin et al. (2011) panel dataset and the CHNS dataset, we have to make sure that the time periods that the two datasets span are as closely overlapping as possible. Thus, in this section we will report sys-GMM estimates for the CHNS dataset during the time period 1989-2004. We will compare these estimates to the sys-GMM estimates that we obtain when using the Benjamin et al. (2011) panel dataset.[1]

In the Benjamin et al. (2011) panel dataset there are two measures of inequality: the mean log deviation, and the 90-10 ratio. We will discuss results for both measures, noting that, in Benjamin et al. (2011) the main measure of inequality is the mean log deviation.

Sys-GMM Estimates of the Interaction Model

For both the Benjamin et al. panel dataset and the CHNS dataset during the period 1989-2004, sys-GMM estimates of the interaction model yield insignificant coefficients on inequality and on the interaction between inequality and initial average income. This can be seen from Appendix Tables 4A and 4B. Appendix Table 4A reports the results for the Benjamin et al. (2011) panel dataset. One can see from Appendix Table 4A that, regardless of whether the dependent variable is the log of the mean of household income per capita (columns (1) and (2)), or the mean of the log of household income per capita (columns (3) and (4)), that the estimates on inequality and the interaction between inequality and initial income per capita are not significantly different from zero at the conventional significance levels. This is the case for both measures of inequality that are available in the published Benjamin et al. (2011) panel dataset, i.e. the mean log deviation (see columns (1) and (3)) and the 90-10 ratio (see columns (2) and (4)).

Appendix Table 4B reports estimates of the interaction model for the CHNS dataset during the period 1989-2004. Appendix Table 4B is structured in exactly the same way as Appendix Table

1 No survey was conducted by the CHNS for the year 2002. There was a survey conducted by the CHNS in the years 2000 and 2004. 1989 is the earliest survey year of the CHNS. We chose 1989-2004 for comparison purposes to the Benjamin et al. panel dataset, which spans the period 1987-2002. Similar results to the ones discussed here for the CHNS data for the period 1989-2004 are obtained for the CHNS dataset during the period 1989-2000.

4A. One can see that in Appendix Table 4B the sys-GMM estimates on inequality and on the interaction between inequality and initial income per capita are not significantly different from zero. This is the case for all four columns of Appendix Table 4B.

Why are the Estimates of the Interaction Model in Appendix Table 4A and 4B Insignificant?

In this sub-section we provide an explanation for why the estimates of the interaction models in Appendix Tables 4A and 4B are not significantly different from zero. The panel models in Appendix Tables 4A and 4B are estimated on a relatively short time-period. In Appendix Tables 4A and 4B the panel spans the late 1980s to early 2000s, i.e. only includes those years when household income per capita was relatively low (compared to the later period 2004-2015). This means that the variance of the variable L.lny, i.e. the log of t-1 household income per capita, that enters the interaction term is not at its maximum.

Our baseline estimates were for a panel of Chinese counties that spanned the period 1989-2015: this is more than one decade longer than the time period of the panel data that the models in Appendix Table 4A and 4B are estimated on. Household incomes per capita have grown substantially in China during the 2000s and 2010s. When estimating the interaction model on the Benjamin et al. panel dataset, or the CHNS data during 1989-2004, one misses the time period when household incomes per capita were relatively high. Using the longest possible time-period when estimating the interaction model ensures that the variance of the time-series variation in initial income is maximized.

The difference in the variance of initial income during the period 1989-2004 and the period 1989-2015 is substantial. According to the CHNS data, the variance of t-1 log household income per capita of Chinese counties during 1989-2015 is more than twice as large as the variance of t-1 log household income per capita of Chinese counties during 1989-2004. During 1989-2004 σ^2(L.lny)=0.19. During 1989-2015 σ^2(L.lny)=0.49.

Sys-GMM Estimates of a Model Without Interaction Term

Appendix Table 4C shows that the Benjamin et al. (2011) panel dataset and the CHNS data for the period 1989-2004 yield similar sys-GMM estimates for a model without an interaction term. Panel A of Appendix Table 4C reports sys-GMM estimates for the Benjamin et al. (2011) dataset. Panel B of Appendix Table 4C shows sys-GMM estimates for the CHNS dataset during the period 1989-2004. Panel A is structured exactly the same as Panel B.

From Appendix Table 4C one can see that, in the Benjamin et al. panel dataset the estimated coefficient on the mean log deviation is positive and significantly different from zero at the 10 percent level or higher. Specifically, in column (1) of Panel A of Appendix Table 4C one can see that for the Benjamin et al. panel dataset the estimated coefficient on the mean log deviation is 0.32 with a standard error of 0.18. The dependent variable in column (1) of Panel A is the log of mean household income per capita. When the dependent variable is the mean of log household income per capita, which is Benjamin et al.'s preferred measure of average income, the estimated coefficient on the mean log deviation is 0.36 and has a standard error of 0.17. Hence, regardless of how average income is measured, either as the log of the mean or the mean of the log, one can see that sys-GMM estimates yield a significant positive average effect of the mean log deviation on income growth in the Benjamin et al. panel dataset.

The same result is obtained for the CHNS data during the period 1989-2004. In column (1) of Panel B of Appendix Table 4C the estimated coefficient on the mean log deviation is 0.12 and has a standard error of 0.06. In column (3) of Panel B of Appendix Table 4C the estimated coefficient on the mean log deviation is 0.39 and has standard error of 0.13.

For the 90-10 ratio, both the Benjamin et al. dataset and the CHNS dataset during 1989-2004 yield sys-GMM estimates that are positive. But these estimates are not significantly different from zero at the conventional significance levels. From column (2) of Panel A of Appendix Table 4C one can see that in the Benjamin et al. dataset, the sys-GMM estimate on the 90-10 ratio is 0.011 and has a standard error of 0.008 when the dependent variable is the log of the mean of household income per capita. When the dependent variable is the mean of the log of household income per capita, see column (4) of Panel A of Appendix Table 4C, the sys-GMM estimate on the 90-10 ratio is 0.009 and has a standard error of 0.007. For the the CHNS dataset during 1989-2004, the sys-GMM estimates on the 90-10 ratio are 0.003 (standard error 0.002) and 0.003 (standard error 0.003), respectively, see columns (2) and (4) of Panel B of Appendix Table 4C.

Comparison to estimates of the interaction model estimates during 1989-2015.

The estimated average effects of inequality on growth for the Benjamin et al. panel dataset are quantitatively and statistically similar to the effect of inequality on growth that can be computed from the estimates of the interaction model during the period 1989-2015 when household income per capita is set equal to the mean household income per capita in the Benjamin et al. panel dataset. For the mean log deviation, the relevant estimates of the interaction model are those shown in column (2) of Table 10. For the 90-10 ratio, the relevant estimates of the interaction model are those shown in column (3) of Table 10.

Consider the interaction model estimates in column (2) of Table 10. For an initial household income per capita equal to the sample average in the Benjamin et al. panel dataset (6.35 logs), the interaction model estimates in column (2) of Table 10 yield a marginal effect of the MLD on growth of 0.321. For comparison, in column (1) of Panel A of Appendix Table 4C, the marginal effect of the MLD on growth is 0.319.

Next, consider the interaction model estimates in column (3) of Table 10. For an initial household income per capita equal to the sample average in the Benjamin et al. panel dataset (6.35 logs), the interaction model estimates in column (3) of Table 10 yield an average effect of the 90-10 ratio on growth of 0.007. For comparison, in column (1) of Panel A of Appendix Table 4C, the marginal effect of the 90-10 ratio on growth is 0.011.

Quantitatively, the sys-GMM estimates for the Benjamin et al. panel dataset in Appendix Table 4C can be interpreted as follows. On average, during 1987-2002: a one standard deviation (0.08) increase in the MLD increased per annum growth by 2.6 percentage points; (ii) a one standard deviation (2.82) increase in the 90-10 ratio increased per annum growth by 3.1 percentage points. These average effects are similar in size to the corresponding average effects that can be obtained from the interaction model estimates in Table 10. According to the estimates of the interaction model in columns (2) and (3) of Table 10, that are based on the CHNS data which spans the longest possible time period, i.e. 1989-2015, setting initial household income per capita equal to the mean in the Benjamin et al. panel dataset the effects of inequality on growth are as follows: (i) a one standard deviation (0.08) increase in the MLD increased per annum growth by 2.7 percentage points; (ii) a one standard deviation (2.82) increase in the 90-10 ratio increased per annum growth by 2.0 percentage points.

Summary: Comparison of Estimates for the CHNS and Benjamin et al. panel dataset

For the sys-GMM estimates of the interaction model on the CHNS dataset spanning the period 1989-2015 one can compute the effect of inequality on growth for an initial income per capita set equal to the mean of income per capita in the Benjamin et al. panel dataset. This effect can be compared to the marginal effect of a model without interaction term that is estimated by sys-GMM on the Benjamin et al. panel dataset. A substantial difference in these effects would suggest that different datasets yield different results. That is not the case. As discussed in detail in the paragraphs above, for an interaction model estimated by sys-GMM on the CHNS dataset during 1989-2015 the estimated effect of inequality on growth for an initial income per capita set equal to the mean of income per capita in the Benjamin et al. panel dataset is similar to the effect of inequality on growth that is obtained by sys-GMM estimation of a model without interaction term on the Benjamin et al.

panel dataset.

A separate issue is with regard to the efficient estimation of the interaction model. The theoretical model of Galor and Zeira (1993) suggests that the effect of inequality on growth differs depending on initial income. To test this hypothesis, which is derived from the theoretical model of Galor and Zeira (1993), an econometric model is required that includes an interaction between initial income and inequality. Efficiently estimating such an interaction model necessitates the use of the maximum amount of time-series data on initial income. The CHNS dataset that spans the period 1989-2015 is optimal with regard to the purpose of efficiently estimating the interaction model: this dataset provides the longest time-series on initial income per capita. The Benjamin et al. panel dataset that spans the period 1987-2002 is not optimal for the purpose of efficiently estimating a panel data model that includes an interaction term between inequality and initial income.

Appendix Table 1: OLS Estimates (Dependent Variable: annualized growth rate)

Inequality	(1) gini	(2) mld	(3) 90-10	(4) 75-25	(5) Q1	(6) Q2	(7) Q3	(8) Q4	(9) Q5
Pool OLS:									
$\text{Inequality}_{i,t-1}$	2.139^{***}	0.955^{***}	0.016	0.138^{***}	-5.168^{**}	-7.882^{***}	-7.665^{**}	2.373	2.187^{**}
	(0.780)	(0.326)	(0.010)	(0.050)	(2.419)	(2.811)	(2.948)	(2.629)	(0.899)
$\text{Inequality}_{i,t-1}$	-0.257^{***}	-0.114^{***}	-0.002	-0.017^{***}	0.589^{**}	0.944^{***}	0.921^{***}	-0.189	-0.265^{***}
$\times\ \text{lny}_{i,t-1}$	(0.085)	(0.037)	(0.001)	(0.006)	(0.274)	(0.308)	(0.323)	(0.285)	(0.099)
$\text{lny}_{i,t-1}$	0.036	-0.021	-0.045^{**}	-0.017	-0.085^{***}	-0.168^{***}	-0.212^{***}	-0.014	0.054
	(0.036)	(0.016)	(0.018)	(0.018)	(0.021)	(0.038)	(0.051)	(0.065)	(0.048)
Observations	474	474	474	474	474	474	474	474	474
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fe model:									
$\text{Inequality}_{i,t-1}$	1.451^{**}	0.492	0.013^{*}	0.108^{*}	-5.447^{**}	-5.977^{**}	-1.944	3.590	1.233
	(0.639)	(0.312)	(0.007)	(0.060)	(2.419)	(2.432)	(2.414)	(2.163)	(0.745)
$\text{Inequality}_{i,t-1}$	-0.164^{**}	-0.058	-0.001^{*}	-0.013^{*}	0.617^{**}	0.676^{**}	0.211	-0.390	-0.139^{*}
$\times\ \text{lny}_{i,t-1}$	(0.073)	(0.037)	(0.001)	(0.007)	(0.273)	(0.286)	(0.273)	(0.242)	(0.083)
$\text{lny}_{i,t-1}$	-0.221^{***}	-0.269^{***}	-0.281^{***}	-0.259^{***}	-0.319^{***}	-0.359^{***}	-0.328^{***}	-0.206^{***}	-0.227^{***}
	(0.040)	(0.028)	(0.024)	(0.030)	(0.029)	(0.037)	(0.040)	(0.051)	(0.052)
Observations	474	474	474	474	474	474	474	474	474
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note. Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level.

† X refers to $\text{Inequality}_{i,t-1}$, $\text{Inequality}_{i,t-1}\times \text{lny}_{i,t-1}$, and $\text{lny}_{i,t-1}$.

Appendix Table 2: Sys-GMM Estimates of a Model Without an Interaction Term (Dependent Variable: annualized growth rate)

	Twostep Sys-GMM			Onestep Sys-GMM		
	(1) baseline	(2) winsor(1%)	(3) limitedlags	(4) baseline	(5) winsor(1%)	(6) limitedlags
$\text{gini}_{i,t-1}$	−0.016	−0.016	−0.079	−0.025	−0.027	−0.046
	(0.088)	(0.099)	(0.100)	(0.074)	(0.082)	(0.076)
$\text{lny}_{i,t-1}$	−0.234***	−0.236***	−0.250***	−0.212***	−0.212***	−0.210***
	(0.043)	(0.043)	(0.032)	(0.027)	(0.027)	(0.027)
AR(2)	0.559	0.536	0.537	0.866	0.864	0.874
p-value	0.576	0.592	0.592	0.387	0.388	0.382
Hansen test	26.042	25.900	9.429	20.819	20.756	13.064
p-value	0.053	0.056	0.151	0.186	0.188	0.110
Instruments:						
No. of IVs	27	27	17	27	27	19
IVs for transformed equation	L(0/7).(X)[†]	L(0/7).(X)	L(4/6).(X)	L(0/7).(X)	L(0/7).(X)	L(0/3).(X)
IVs for level equation	D.X	D.X	D.X	D.X	D.X	D.X
Underidentification test:						
Kleibergen-Paap LM stat	38.848	38.688	27.547	38.848	38.688	36.639
p-value	0.002	0.002	0.000	0.002	0.002	0.000
Weak identification test:						
Kleibergen-Paap Wald stat	18.369	17.993	11.993	18.369	17.993	26.666
Stock-Yogo critical values:						
10% maximal IV relative bias	11.00	11.00	10.22	11.00	11.00	10.58
20% maximal IV relative bias	6.14	6.14	6.20	6.14	6.14	6.23
30% maximal IV relative bias	4.43	4.43	4.73	4.43	4.43	4.66
Observations	474	474	474	474	474	474
Number of county	72	72	72	72	72	72
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Note. Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level.

[†] X refers to $\text{gini}_{i,t-1}$ and $\text{lny}_{i,t-1}$.

Appendix Table 3: Controlling for Inequality Squared

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inequality	gini	mld	90-10	75-25	Q1	Q2	Q3	Q4	Q5
$\text{Inequality}_{i,t-1}$	2.421**	1.147***	0.027**	0.286***	−9.147	−5.618	−5.945*	6.941	1.876
	(1.170)	(0.340)	(0.013)	(0.100)	(5.688)	(5.092)	(3.543)	(8.048)	(1.416)
${\text{Inequality}_{i,t-1}}^2$	−0.155	0.029	0.000	−0.003**	1.022	−6.110	1.374	−1.794	−0.195
	(0.266)	(0.064)	(0.000)	(0.001)	(1.042)	(7.319)	(6.857)	(8.522)	(0.259)
$\text{Inequality}_{i,t-1} \times \text{lny}_{i,t-1}$	−0.267**	−0.134***	−0.003**	−0.030***	1.093	0.860*	0.656*	−0.627	−0.195
	(0.133)	(0.037)	(0.002)	(0.011)	(0.663)	(0.500)	(0.329)	(0.554)	(0.158)
$\text{lny}_{i,t-1}$	−0.076	−0.169***	−0.207***	−0.119**	−0.214***	−0.262***	−0.288***	−0.052	−0.099
	(0.074)	(0.034)	(0.032)	(0.052)	(0.050)	(0.063)	(0.058)	(0.133)	(0.087)
AR(2) test	0.385	0.091	0.446	0.688	0.736	0.640	0.556	0.533	0.499
p-value	0.700	0.928	0.656	0.491	0.462	0.522	0.578	0.594	0.618
Hansen test	37.426	32.974	33.155	35.210	39.567	36.177	34.692	37.294	36.194
p-value	0.234	0.419	0.411	0.319	0.168	0.280	0.341	0.239	0.279
Instruments:									
No. of IVs	45	45	45	45	45	45	45	45	45
IVs for transformed equation	L(0/7).(X)†	L(0/7).(X)	L(0/7).(X)	L(0/7).(X)	L(0/7).(X)	L(0/7).(X)	L(0/7).(X)	L(0/7).(X)	L(0/7).(X)
IVs for level equation	D.X	D.X	D.X	D.X	D.X	D.X	D.X	D.X	D.X
Underidentification test:									
Kleibergen-Paap LM stat	48.552	52.160	51.745	47.560	48.634	48.501	51.470	49.736	50.994
p-value	0.040	0.018	0.020	0.048	0.039	0.040	0.021	0.031	0.024
Weak identification test:									
Kleibergen-Paap Wald stat	43.105	37.175	110.896	56.661	12.004	60.201	66.679	131.654	33.041
Observations	474	474	474	474	474	474	474	474	474
Number of county	72	72	72	72	72	72	72	72	72
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note. † X refers to $\text{Inequality}_{i,t-1}$, ${\text{Inequality}_{i,t-1}}^2$, $\text{Inequality}_{i,t-1} \times \text{lny}_{i,t-1}$, and $\text{lny}_{i,t-1}$. The method of estimation is system-GMM.

Appendix Table 4A: Sys-GMM Estimates of the Interaction Model
Benjamin et al. (2011) Panel Dataset, Time Period: 1987-2002

	(1)	(2)	(3)	(4)
lny	log mean income per capita		mean log income per capita	
Inequality	mld	90-10	mld	90-10
$\text{Inequality}_{i,t-1}$	-0.970	-0.073	-0.476	-0.065
	(2.420)	(0.125)	(2.142)	(0.117)
$\text{Inequality}_{i,t-1} \times \text{lny}_{i,t-1}$	0.198	0.013	0.135	0.012
	(0.375)	(0.020)	(0.336)	(0.018)
$\text{lny}_{i,t-1}$	-0.146*	-0.153	-0.143*	-0.151
	(0.082)	(0.112)	(0.073)	(0.104)
AR(2) test	-0.549	-0.533	-0.195	-0.186
p-value	0.583	0.594	0.846	0.852
Hansen test	5.707	6.342	4.657	5.553
p-value	0.457	0.386	0.589	0.475
Instruments:				
No. of IVs	15	15	15	15
IVs for differenced equation	L(1/2).(X)†	L(1/2).(X)	L(1/2).(X)	L(1/2).(X)
IVs for level equation	D.X	D.X	D.X	D.X
Underidentification test:				
Kleibergen-Paap LM stat	43.761	39.593	44.171	40.524
p-value	0.000	0.000	0.000	0.000
Weak identification test:				
Kleibergen-Paap Wald stat	14.190	12.901	14.758	12.598
Stock-Yogo critical values:				
10% maximal IV relative bias	9.37	9.37	9.37	9.37
20% maximal IV relative bias	5.78	5.78	5.78	5.78
30% maximal IV relative bias	4.46	4.46	4.46	4.46
Observations	490	490	490	490
Number of county	82	82	82	82
Time FE	Yes	Yes	Yes	Yes

Note. † X refers to $\text{Inequality}_{i,t-1}$, $\text{Inequality}_{i,t-1} \times \text{lny}_{i,t-1}$, and $\text{lny}_{i,t-1}$.

Appendix Table 4B: Sys-GMM Estimates of the Interaction Model
CHNS Dataset, Time Period: 1989-2004

Iny	(1) log mean income per capita	(2)	(3) mean log income per capita	(4)
Inequality	mld	90-10	mld	90-10
$\text{Inequality}_{i,t-1}$	0.128	−0.021	1.452	−0.032
	(0.992)	(0.023)	(1.164)	(0.026)
$\text{Inequality}_{i,t-1} \times \text{lny}_{i,t-1}$	−0.005	0.003	−0.145	0.004
	(0.120)	(0.003)	(0.149)	(0.003)
$\text{lny}_{i,t-1}$	−0.129*	−0.250***	−0.057	−0.310***
	(0.076)	(0.082)	(0.087)	(0.071)
AR(2) test	0.895	0.987	1.325	1.487
p-value	0.371	0.324	0.185	0.137
Hansen test	9.965	10.360	18.504	19.163
p-value	0.126	0.110	0.101	0.085
Instruments:				
No. of IVs	10	10	16	16
IVs for transformed equation	L(0/1).(X)†	L(0/1).(X)	L(0/3).(X)	L(0/3).(X)
IVs for level equation	D.X	D.X	D.X	D.X
Underidentification test:				
Kleibergen-Paap LM stat	31.367	31.844	33.632	32.895
p-value	0.000	0.000	0.001	0.002
Weak identification test:				
Kleibergen-Paap Wald stat	37.303	52.968	43.505	52.531
Stock-Yogo critical values:				
10% maximal IV relative bias	9.37	9.37	10.33	10.33
20% maximal IV relative bias	5.78	5.78	5.94	5.94
30% maximal IV relative bias	4.46	4.46	4.37	4.37
Observations	240	240	240	240
Number of county	54	54	54	54
Time FE	Yes	Yes	Yes	Yes

Note. † X refers to $\text{Inequality}_{i,t-1}$, $\text{Inequality}_{i,t-1} \times \text{lny}_{i,t-1}$, and $\text{lny}_{i,t-1}$.

Appendix Table 4C: Sys-GMM Estimates of a Model Without Interaction Term

Iny	(1) log mean income per capita	(2)	(3) mean log income per capita	(4)
Inequality	mld	90-10	mld	90-10
Panel A: Benjamin et al. (2011) panel dataset, time period 1987-2002				
Inequality$_{i,t-1}$	0.319*	0.011	0.355**	0.009
	(0.181)	(0.008)	(0.169)	(0.007)
lny$_{i,t-1}$	−0.146***	−0.140***	−0.156***	−0.143***
	(0.045)	(0.043)	(0.044)	(0.043)
p-value of AR(2) test	0.647	0.729	0.893	0.975
p-value of Hansen test	0.459	0.529	0.622	0.563
p-value of Kleibergen-Paap LM stat	0.000	0.000	0.000	0.000
Kleibergen-Paap Wald F stat	100.635	88.030	102.906	99.914
Observations	490	490	490	490
Number of county	82	82	82	82
Time FE	Yes	Yes	Yes	Yes
Panel B: CHNS Dataset, time period 1989-2004				
Inequality$_{i,t-1}$	0.118**	0.003	0.395***	0.003
	(0.057)	(0.002)	(0.131)	(0.003)
lny$_{i,t-1}$	−0.161***	−0.186***	−0.176***	−0.226***
	(0.060)	(0.059)	(0.054)	(0.058)
p-value of AR(2) test	0.436	0.367	0.558	0.147
p-value of Hansen test	0.149	0.114	0.503	0.152
p-value of Kleibergen-Paap LM stat	0.000	0.000	0.001	0.000
Kleibergen-Paap Wald F stat	27.163	60.588	10.358	11.559
Observations	240	240	240	240
Number of county	54	54	54	54
Time FE	Yes	Yes	Yes	Yes

Note. The method of estimation is system-GMM. In Panel A, Huber robust standard errors (shown in parentheses) are clustered at the village level; in Panel B, Huber robust standard errors are clustered at the county level.