SI Appendix

Materials and Methods

Data Sources

Our main data source is the 2015 China Rural Household Panel Survey (CRHPS, 2015) conducted by Zhejiang University. It was used to establish the relationship between agricultural chemical use intensities and farm size. The CRHPS is a nationally representative survey covering all provinces except Xinjiang and Tibet. The original rural household data include 22,535 households that are registered as agricultural residents. These households consist of 76,675 individuals from 1,439 residential committees and villages, located in 363 selected counties in China (Fig. S6). Because the survey reported only the sowing area and yield of six major crops (rice, wheat, maize, bean, peanut and rapeseed), we focus on the households that cultivated those major crops only in our main analysis. The CRHPS data are open to all researchers free of charge and full access to all data, after they register an account at the Zhejiang University's server with following steps:

1. Download the registration form from http://ssec.zju.edu.cn/dataset/CRHPS/ and fill in his/her information;
2. Send the filled registration form to rwskdata@zju.edu.cn;
3. Confirmation from the staff of the CRHPS data server with a username and password to access the database and other instructions on how to use the database.
4. Login the database with full accessing right to all data.

The survey collected information on household demographic features, agricultural and non-agricultural activities, and household income from these activities. It also collected information on the residential committees and villages where the households resided. A detailed run-down of all the variables used in our paper is provided in a later section.

We used two other independently collected data sources to check the robustness of our results.

(1) China Household Income Project (CHIP, 2002). The CHIP is available at http://www.ciidbnu.org/chip/index.asp hosted at Beijing Normal University. It is a widely used nationally representative household survey since 1988. We used farm size and agricultural intermediate inputs expenditure (including agricultural chemicals and other inputs[1]) for each household to test the robustness of the negative intensity-size relationship we established using the CRHPS.

(2) The Second National Agricultural Census (NAC, 2006). The NAC was conducted by the National Bureau of Statistics of China. It is a decennial census that collects information of a nationally representative sample of rural households, as well as village and township governments, etc. We used the physical quantities of chemical fertilizer consumption aggregated according to farm size group. The data are from the *Data Collection of the Second National Agricultural Census in China*. The pattern we established from the NAC is consistent with that from the CRHPS, providing evidence in support of the data quality of the CRHPS (Fig. S6).

We also used the following data sources for the analysis on agricultural chemical use and farm size.

(3) The FAO (Food and Agriculture Organization) database of the United Nations. Most of the variables used for the international comparisons in the paper were compiled from the FAO database. The data are available for downloading at http://www.fao.org/statistics/en/. The

[1] Unfortunately, we are not able to identify what exactly is included in those "other inputs" due to data limitations.

www.pnas.org/cgi/doi/10.1073/pnas.1806645115

variables we used include the annual consumption and prices of chemical fertilizers, areas of arable land and permanent crops, harvest areas, yields, outputs and prices of crops, etc.

(4) The World Bank Open Data. The variables we used are country-level GDP per capita (in current US$), purchasing power parity adjusted GDP per capita (in current international $), value-added of agriculture (% of GDP), arable land per capita, and proportion of urban population from 1961 to 2010. All the variables were compiled from the World Bank database, available at http://data.worldbank.org/.

(5) Various National Statistical Reports. We used the *Annual Report of China Agricultural Development* to compile the number of residents per household, the per capita cultivated area, and the price index of fertilizer in China. We constructed the longitudinal data of the price index of cereal crops from the *Yearbook of China Agricultural Product Price*. We used the country-level longitudinal data of average household arable land compiled by Lowder et al. (1) from decennial national agricultural censuses during 1960-2000. Data for nitrogen fertilizer use for cereals (rice, wheat and maize) are from Zhang et al. (2).

A Detailed Introduction to the CRHPS

The Sampling Design

The CRHPS was initiated by the Southwestern University of Finance and Economics in 2011 (3). The survey was since then carried out every two years. The Social Science Research Center of Zhejiang University joined and led the CRHPS since 2015 (the third round), and focused more on the agriculture sector, rural areas, and farmers. The CRHPS employs a stratified three-stage probability proportion to size (PPS) random sample design (4). The primary sampling units are counties (including county-level cities and districts) from all provinces in China except Tibet, Xinjiang, Hong Kong, Macau, and Taiwan. The sampling practice is performed with the PPS method and the samples are weighted by their gross domestic products (GDP). The second stage of sampling selects the residential committees/villages from the counties/cities drawn in the first stage. The samples are weighted by their population size. The last stage of sampling selects households by systematic sampling from the residential committees/villages from the second stage. Our previous work explains the sampling method in more details (3).

The Questionnaire

The questionnaire has two modules: rural and urban. Depending on whether a respondent is registered as a rural or urban resident, the corresponding module is invoked during the interview. Each module contains four sections and respondents are asked to answer the questions based on the information of the previous 12 months (which is the reference period of the survey). The structure of the questionnaire is provided below. The complete questionnaire can be accessed at the data portal: http://ssec.zju.edu.cn.

I. Household's Demographic Characteristics
- a. Household demographics.
- b. The jobs and incomes of family members
- c. Financial knowledge and the evaluation of local governments

II. Assets and Debts
- a. Financial and non-financial assets
- b. Household debts
- c. Financial fraudulence

III. Social and Commercial Insurance
- a. Social insurance
- b. Commercial insurance

IV. Expenditure and Income

a. Consumption
b. Transfers
c. Others

Training the Interviewers
Owing to the large size of the sample, over 1,000 interviewers were recruited to conduct the interviews. To ensure the interview quality, all interviewers received trainings on interviewing skills, the content of the questionnaire, and the use of the CAPI (Computer-assisted Personal Interviewing) system. The training taught the interviewers how to collect precise information from the household. In particular, the interviewers received trainings on how to identify and build trust with qualified interviewees, how to ask questions politely and accurately, and how to interpret the questionnaire correctly, etc. The interviewers were given a pad-computer pre-installed with both the CAPI system and the survey management system. All interviewers went through practice interviews before being dispatched to the field.

Interview Quality Control
The CAPI system helps control the quality of the interview in several ways. First, there are strict rules for replacing samples unavailable during the first or second attempt. Usually a sampled case is replaced only after at least three failed attempts. Refusal rate varies from 2011 to 2015, but never exceeds 4% thanks to the assistance of the local communities. Second, the CAPI ensures that the logical relationship between questions are always respected. For instance, for a household with no livestock raised, if the interviewer enters a strictly positive income from livestock, the CAPI will notify the interviewer about the error and refuse to proceed ahead. The CAPI also requires that data inputs for many questions be constrained within realistic ranges. The system will notify the user when unrealistic data have been entered. This reduces the occurrence of unintended errors substantially.

Data Quality Control
The data were transferred back to the university upon the completion of an interview. The quality of the raw data were checked by multiple ways. First, the computer conducted simple statistical analysis of the original audio records, and checked the contact information of the sample to ensure that there are no obvious faking samples. Suspicious samples were flagged and the original audio record later were manually investigated. Second, the raw audio records of a proportion of randomly selected samples were listened again to check for errors or faking. Third, all the households were called to verify that the interview was conducted appropriately after the completion of the survey. Fourth, several aggregate statistical moments from the data such as income and demographics were validated by comparing with the corresponding moments calculated from national censuses conducted by the National Statistics Bureau of China. Some of these results were reported in our previous publications. See Gan et al. (3) for more details.

Variables from the CRHPS 2015

Agricultural Chemicals Use Intensity

Throughout the paper, agricultural chemicals refer to chemical fertilizers and pesticides. Unless otherwise specified, agricultural chemical use intensity is defined as household's expenditure on chemical fertilizers and/or pesticides per sowing area.

Sowing Area and Contractual Area

Sowing area refers to household's gross cultivated land areas of several major crops (rice, wheat, maize, bean, peanut and rapeseed) in the four seasons immediately preceding the survey.

We use the term *sowing area* as our main measure of *farm size* in China in the main text. Household's contractual area refers to the area of tenured arable land according to the rural Household Contract Responsibility System (HCRS).

Plant Type

Plant type is a categorical variable denoting the type of crops that each household grows. We focus on the major grain crops (rice, wheat and maize) in our main results, but test for the robustness of our results by further including the major cash crops (bean, peanut and rapeseed).

Multiple Land Parcels

Multiple Land Parcels is a dummy variable that equals one if the household has more than one parcel of arable land. Under the HCRS, farm land is distributed based on an egalitarian principle. To ensure that everyone gets land parcels that are on average of the same quality, hence not all land parcels are contingent to each other.

Land Type and Land Quality

Both variables are categorical variables. Land type takes four values: paddy land, irrigated land, upland, and others. Land quality is a subjective evaluation by the farmer on the fertility of the household's largest parcel, with a five-point Likert scale.

Terrain

Terrain is also a categorical variable describing six different geographical terrains of a household's location: plains, mountains, seaside, basin, plateau, and hilly area.

Region

Region is a categorical variable including six province groups: north China (Beijing, Tianjin, Shanxi, Hebei, and Inner Mongolia), northeast China (Heilongjiang, Jilin, and Liaoning), east China (Anhui, Fujian, Jiangxi, Zhejiang, Shanghai, and Shandong), south and central China (Henan, Guangdong, Guangxi, Hainan, Hubei, and Hunan), southwest China (Sichuan, Yunnan, Chongqing, and Guizhou), and northwest China (Shaanxi, Qinghai, Ningxia, and Gansu). The 2015 CRHPS does not include Tibet and Xinjiang.

Agricultural Output per ha (Land Productivity)

Land productivity is the household agricultural output divided by the total sowing area. Agricultural output is reported by the interviewees. To test the accuracy of the reported value of output, we also calculated the agricultural output by multiplying the physical quantity of crop yields with regional prices of crops. Regional prices are imputed from households who report both the quantity of crop yields and their market values. Reassuringly, the reported and imputed output values are highly in accordance with each other.

Agricultural Output per Labor (Labor Productivity)

Labor productivity is the household agricultural output divided by the household agricultural labor input. Agricultural labor input refers to working hours from both the household members working on family farms and hired agricultural workers.

Agricultural Income per Labor

Agricultural income per labor is calculated as household agricultural value-added divided by the household agricultural labor input. Household agricultural value-added is calculated as household agricultural output value less agricultural costs, which include expenditure on seeds, chemical fertilizers, pesticides, energy and water, and the per-period equivalent cost of investment on agricultural machinery. The per-period cost of capital equals 10% depreciation rate plus a 6% interest rate on the original value. The 10% is a widely used annual depreciation rate for China, and the 6% is the median interest rate obtained from the household data of the CRHPS 2015.

Non-agricultural Income per Labor (The Opportunity Cost of Farming)

Non-agricultural income per labor is defined as the rural household's expected wage rate if they move to the non-agricultural sectors conditional on their education, working experience, etc. The expected wage rate was estimated using the *Mincerian* Earnings Equation. More details on the construction of this variable can be found in the section "*Mincerian* Earnings Equation and the Opportunity Cost of Farming."

Rent

Land rent is the median per ha value in the CRHPS 2015. We excluded the bottom 5% of land rent to exclude token charges between relatives and acquaintances. The estimated land rent is 4,500 Yuan (RMB) per ha.

N-losses

N-losses (nitrogen surplus) is used as an indicator of environment damages. It is calculated as the difference between the amounts of nitrogen applied to and those eventually contained in the harvested plants. The application rate was imputed by assuming that the average nitrogen application rate in the sample equals those in Chen et al. (5). Application rates by size group were then calculated by comparing the group average expenditure with the overall average. The amounts of nitrogen content contained in harvested crops (N_{har}) were calculated using crop output data from the CRHPS and the absorbing rate from the literature (6, 7). The N-losses was then calculated as:

$$\mathrm{N}_{loss} = \sum_{i} \frac{(N_{app}^{i} - N_{har}^{i}) \cdot 100\%}{N_{app}^{i}}$$

where subscript i stands for different crops.

Variables from Other Data Sources

Adjusted Average Household Farm Size

The adjusted average household farm size in Fig. 2 for each country is measured as the residual from a regression of average farm size on arable land per capita. The reason for this adjustment is that the average farm size in a country could be large simply because it is more land-abundant. By excluding the difference in land endowment, we are left with the variations in farm size that associate with agricultural production structure.

Fertilizer-crop Price Ratio

Fertilizer-crop price ratio is measured by the price of urea (P_{urea}) over the weighted price of rice (P_{rice}), wheat (P_{wheat}), and maize (P_{maize}) (with their harvest area as weights):

$$price\ \ ratio = \frac{P_{urea}}{(P_{rice} \cdot area_{rice} + P_{wheat} \cdot area_{wheat} + P_{maize} \cdot area_{maize}) / area_{rice,wheat,maize}}$$

Percentage of Land Growing Legumes and Vegetables

Annual Legumes% and vegetables% are a country's harvested area of leguminous plants and vegetable and fruits over its total harvested area.

Sowing Area, Contractual Area, Agricultural Intermediate Input

These variables were taken from the CHIP 2002. The definitions of sowing area and contractual area are the same as in the CRHPS. Agricultural Intermediate Input are materials and services used for agricultural production, including chemicals, machinery, seeds, irrigation, labor, rend, etc.

Fertilizer and Pesticide Use by Farm Size

The fertilizer and pesticide consumption by farm size consumption were compiled from the NAC (2006). They were calculated by dividing the group aggregate fertilizer and pesticide consumption over the group aggregate land area.

Agricultural Productivity

The agricultural productivity is obtained using agricultural value-added divided by rural population. And agricultural value-added is obtained by GDP multiplying share of agricultural value-added over GDP. Data used to calculate agricultural productivity are from the World Bank database.

Temperature and precipitation

National level temperature and precipitation data during 1961-2010 are obtained from the Centre for Environmental Data Analysis (http://www.ceda.ac.uk).

Method

Cross-Sectional Analysis of Agricultural Chemical Use Intensity and Land Productivity with Farm Size in China

We used the CRHPS to estimate the relationship between agricultural chemical use intensity and farm size, controlling for a series of confounding factors. The regression equation is:

$$Agricultural\_Chemical\_use_i = \alpha + \gamma \cdot farmsize_i + \sum_k \beta_k x_{ki} + \varepsilon_i, \quad (1)$$

where subscript i denotes households. $Agricultural\_Chemical\_use$ is the logarithm of agricultural chemicals (chemical fertilizers and/or pesticides) use intensity. $farmsize$ refers to the logarithm of sowing area. We also used the logarithm of contractual area to check the robustness of our results. The confounding factors that we controlled for include dummy variables for multiple land parcels, plant type, terrain, land type and quality.

To reduce measurement errors, we excluded samples that spent only 1 yuan (RMB) on chemical inputs and those whose yield was beyond the limit that can be achieved by the current state of the art technology. The latter is 1500 kg/mu, 1000 kg/mu and 2000 kg/mu for rice, wheat, and corn respectively. We further excluded data for the top 0.5% sample of the three chemical inputs, as well as a suspicious outliner such as who claimed 4,000 ha of sown land area whilst only owning 2.67 ha of contracted land area. The estimated results were reported in Table S2.

By replacing the agricultural chemicals use intensity with land productivity, we estimated the relationship between land productivity and farm size:

$$output_i = \alpha + \gamma \cdot farmsize_i + \sum_k \beta_k x_{ki} + u_i. \ (2)$$

The results were reported in the last two columns of Table S3.

Longitudinal Analysis of International Agricultural Chemical Use Intensity and Land Productivity with Farm Size

To check the relationship between agricultural chemical use per land area and farm size, we estimated a fixed effect model using data from the FAOSTAT and the World Bank database:

$$Fertilizer\_use_{jt} = \alpha + \gamma \cdot farmsize_{jt} + \sum_m \varphi_m z_{mjt} + \sigma_j + \mu_{jt} \quad (3)$$

where subscript j and t denotes country and time respectively. The agricultural chemical application rate $Fertilizer\_use_{jt}$ for each country was calculated as the average total fertilizer use per land area (arable land and permanent crops) for each decade between the 1960s and 2000s using the FAOSTAT. $farmsize$ is the logarithm of the average size of agricultural households for each decade. z_m's include logarithm of GDP per capita, logarithm of fertilizer-crop price ratio, and crops mix proxied with country j's legumes% and vegetables%. The regression results were reported in the first four columns of Table 2.

By replacing the agricultural chemicals use intensity with land productivity, we built the relationship between farm size and land productivity:

$$crop\_yield_{it} = \alpha + \gamma \cdot farmsize_{jt} + \sum_m \varphi_m z_{mjt} + \sigma_j + \tau_{jt}, \quad (4)$$

where $crop\_yield_{it}$ is the production of cereals (rice milled equivalent) per arable land area. The estimated results were also reported in the last column of Table 2.

<u>*Mincerian* Earnings Equation and the Opportunity Cost of Farming</u>

We used the classical *Mincerian* Equation to calculate the opportunity cost of being a farmer. The opportunity cost is defined as the expected income of an individual currently working on a family farm if s/he migrates to the non-agriculture sector in the urban area. The expected income was then calculated as the fitted value from the *Mincerian* regression. To do this, we first estimated the following *Mincerian* Equation using the CRHPS data of the six most popular non-agricultural industries for rural migrants:

$$\ln w = \ln w_0 + \rho \cdot s + \theta_1 \cdot Expe + \theta_2 \cdot Expe^2$$
$$+\delta_1 \cdot age + \delta_2 \cdot age^2 + \delta_3 \cdot urban + \epsilon. \quad (5)$$

The six industries are construction, manufacturing, housekeeping and repair services, food and restaurant services, agriculture management and services, and other temporary or seasonal jobs. According to the CRHPS, these six sectors together employed more than 66% of rural migrants. In the regression, $\ln w$ is the logarithm of individual earnings, s is the years of schooling, $Expe$ is years of working experience in a current job, age is an individual's age, and $urban$ is a dummy variable that equals 1 if the individual lives in the urban area.

With the estimates from the regression, we then calculated the potential non-agricultural earnings in urban areas for each farmer and the potential non-agricultural earnings for the family using the fitted values from the equation based on household's demographic characteristics. Specifically, we set $\text{Expe} = 0$ for all households. We then added land rent to the fitted values, and further reduce the fitted values by 30% to account for the higher living costs in urban areas (8).

We did not control for industries and provinces in the *Mincerian* regression since we cannot predict in which industry and province the migrated workers will work in. However, we do find that the residuals from the *Mincerian* regression are on average symmetric around zero, suggesting that when the Law of Large Number is invoked, on average the predicted value within farm size group would be similar.

Scenario Analysis

We conducted a series of scenario analyses to study how the agricultural chemical use, fertilizer loss (accounted as N), output, and household income would change if we changed the farm size distribution in China in several ways. First, we removed the policy distortions mentioned in our main text, so farmers with sowing area smaller than 1.1 ha would move to non-agriculture occupations and rent their farmland to the group with >1.1 ha. Second, we increased China's average farm size in 2010 to the level predicted by the fitted line in Fig. 3A by reallocating the land of small farms to large farms. In the third analysis, we again reallocated the land of small farmers to large farmers, but to the world average of 6.1 ha.

We created grids of farms with different sizes, $s_1,\ldots,s_N$. For farms of size $s_i, i \in \{1,\cdots,N\}$, we constructed the within-group average of the following variables as inputs to our scenario analysis:

1. Share of farms, i.e., the share of total number of farms in each size group: $\theta_1,...,\theta_N$
2. Share of farmland, i.e., the share of total land areas in each size group: $\phi_1,...,\phi_N$
3. Crop yield per ha: $y_1,...,y_N$
4. Fertilizer use per ha: $f_1,...,f_N$
5. N losses per ha: $loss_1,...,loss_N$
6. Pesticide use per ha: $p_1,...,p_N$
7. Agricultural income of farmers operating on farm size i: $w_1^a,...,w_N^a$
8. Hypothetical non-agricultural income (including land rents) if farmers moved to urban non-agriculture occupations: $w_1^n,...,w_N^n$

Using these variables as inputs, we computed the following aggregate variables:
1. Average yield per ha: $Y = \sum_{i=1}^{N} \phi_i y_i$
2. Average fertilizer use per ha: $F = \sum_{i=1}^{N} \phi_i f_i$
3. Average N losses per ha: $Loss = \sum_{i=1}^{N} \phi_i loss_i$
4. Average pesticide use per ha: $P = \sum_{i=1}^{N} \phi_i p_i$
5. Average income for farmers: $W = \sum_{i=1}^{N} \theta_i w_i^a$

In the above calculation we assumed that for all variables, all farms in particular groups share the same value of the group average. That is, if there were 100 farms of size 5 mu, we assumed that all these 100 farms had the same fertilizer use per ha, etc. We suggest that this assumption is not distorting because the group size is only 1 Mu (1/15 ha).

We implemented our scenario analyses as follows. First, for each scenario, we found the cut-off size s^*, which is the minimal size of farm after land and labor reallocation. Different scenarios are represented by different s^*. This provides a direct observable measure to map our scenarios analyses to real world policies. In scenario 1, where we eliminated policy distortions, s is the threshold such that $w_i^a \geq w_i^n$ for all $s_i \geq s$. In scenarios 2 and 3, we set s so that after land reallocation, the average farm size equals respectively to the level predicted by regressions and the world average. Second, we moved farmers with farmland $s_i < s^*$ to non-agricultural sectors, and redistributed their farmland to the remaining farmers in proportion to their initial share of farmland ϕ_i. The resulting new distribution of farmland is

$$\phi_i^* = \begin{cases} 0 & s_i < s^* \\ \dfrac{s_i}{\sum_{s_j \geq s^*} s_j} & s_i \geq s^* \end{cases}.$$

We then calculated the remaining farmers' new agricultural income w_i^{a*} from their increased farm size. Finally, we computed the resulting new aggregate variables as follows:

1. Average yield per ha: $Y^* = \sum_{i=1}^{N} \phi_i^* y_i$

2. Average fertilizer use per ha: $F^* = \sum_{i=1}^{N} \phi_i^* f_i$

3. Average N losses per ha: $Loss^* = \sum_{i=1}^{N} \phi_i^* loss_i$

4. Average pesticide use per ha: $P^* = \sum_{i=1}^{N} \phi_i^* p_i$

5. Average income for farmers: $W^* = \sum_{s_i \geq s^*} \theta_i w_i^{a*} + \sum_{s_i < s^*} \theta_i w_i^n$

Table S1 Number of farms in each farm group used for the analysis in Figure 1 and 3

Average group size (ha)	Group range (mu)	Number of farms for analysis			
		Fertilizer	Pesticide	Yield	Labor productivity
0.0	0.1-0.9	76	66	79	75
0.1	1-1.9	237	213	248	245
0.1	2-2.9	312	267	333	329
0.2	3-3.9	301	275	312	308
0.3	4-4.9	279	259	281	279
0.3	5-5.9	212	194	218	215
0.4	6-6.9	223	206	234	231
0.5	7-7.9	126	122	131	130
0.5	8-8.9	164	151	164	163
0.6	9-9.9	66	62	69	68
0.7	10-10.9	223	212	230	229
0.7	11-11.9	53	46	53	51
0.8	12-12.9	102	93	102	101
0.9	13-13.9	53	49	52	52
0.9	14-14.9	58	53	59	58
1.0	15-15.9	58	53	60	60
1.1	16-16.9	49	49	51	50
1.1	17-17.9	16	12	16	16
1.2	18-18.9	38	33	41	41
1.3	19-19.9	9	6	9	9
1.4	20-24.9	178	165	174	173
1.7	25-29.9	47	43	45	44
2.0	30-34.9	69	64	73	73
2.4	35-39.9	16	13	15	15
2.7	40-44.9	35	30	35	34
3.0	45-49.9	15	13	15	15
3.4	50-59.9	35	33	33	33
4.7	60-99.9	60	54	59	59
14.5	>100	59	47	61	61

Note: farm numbers varied across different analysis factors (Fertilizer, Pesticide, Yield, Labor productivity) due to the controls on data quality. 15 mu = 1 ha.

Table S2 The relationship between agricultural chemical use intensity and farm size based on CRHPS 2015 data

	Fertilizer		Pesticide		Fertilizer + Pesticide			
	Grains	Grains	Grains	Grains	Grains	Gross	Grains	Gross
Sowing area	-0.338***	-0.357***	-0.487***	-0.467***	-0.357***	-0.354***	-0.369***	-0.376***
	(0.024)	(0.027)	(0.028)	(0.032)	(0.024)	(0.023)	(0.027)	(0.026)
Land parcels	0.042	0.036	0.022	-0.006	0.055	0.076	0.046	0.064
	(0.052)	(0.052)	(0.066)	(0.065)	(0.052)	(0.053)	(0.052)	(0.053)
Plant type	YES	YES	YES	YES	YES	YES	YES	YES
Land type	YES	YES	YES	YES	YES	YES	YES	YES
Land quality	YES	YES	YES	YES	YES	YES	YES	YES
Terrain	YES	YES	YES	YES	YES	YES	YES	YES
Region	NO	YES	NO	YES	NO	NO	YES	YES
Constant	7.579***	7.284***	6.656***	6.344***	7.895***	7.694***	7.555***	7.277***
	(0.093)	(0.112)	(0.107)	(0.138)	(0.095)	(0.160)	(0.116)	(0.165)
N	3048	3048	2773	2773	3093	3712	3093	3712
F stat	15.53	14.20	44.65	39.46	19.01	17.23	17.36	16.83
adj. R^2	0.191	0.204	0.352	0.368	0.216	0.178	0.232	0.202

Note: Robust standard errors are reported in parentheses, with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data used in Table S2 were from the 2015 China Household Panel Survey (CRHPS, 2015). Columns labeled Grains refer to the group of households that cultivated the three major grain crops (rice, wheat and maize) only, which consist of 3,048 households. Columns labeled Gross refer to the group of rural households that cultivated three major grain crops and/or three major cash crops (bean, peanut and rapeseed). The dependent variable in Table S2 is the logarithm of agricultural chemicals per land area, measured by the household expenditure on chemical fertilizer and/or pesticide per sowing area or contractual area. Detailed explanations of independent variables can be found in the section "Variables from the CRHPS 2015."

Table S3 Agricultural output per ha, Agricultural output per labor, and farm size based on the CRHPS 2015 data

	Output per ha				Output per labor			
	Grains	Grains	Grains	Grains	Grains	Grains	Grains	Grains
Sowing area	-0.017		-0.033		0.869***		0.826***	
	(0.018)		(0.024)		(0.022)		(0.028)	
Contract area		0.034		0.000		0.697***		0.520***
		(0.022)		(0.029)		(0.030)		(0.035)
Fertilizer per Ha	0.074***	0.071**	0.058*	0.068**				
	(0.023)	(0.022)	(0.024)	(0.022)				
Land parcels	-0.021	-0.035	-0.015	-0.035	0.031	0.147*	0.046	0.181**
	(0.041)	(0.041)	(0.042)	(0.041)	(0.051)	(0.060)	(0.051)	(0.058)
Plant type	YES	YES	YES	YES	YES	YES	YES	YES
Land type	YES	YES	YES	YES	YES	YES	YES	YES
Land quality	YES	YES	YES	YES	YES	YES	YES	YES
Terrain	YES	YES	YES	YES	YES	YES	YES	YES
Province	NO	NO	YES	YES	NO	NO	YES	YES
Constant	9.340***	9.984***	9.417***	9.387***	8.950***	8.701***	8.932***	9.005***
	(0.183)	(0.071)	(0.191)	(0.071)	(0.096)	(0.119)	(0.112)	(0.175)
N	2900	2879	2900	2879	3043	3033	3043	3033
F stat	9.91	10.56	6.89	6.34	101.84	48.67	68.27	28.85
adj. R^2	0.096	0.099	0.118	0.119	0.521	0.331	0.532	0.380

Note: Robust standard errors are reported in parentheses, with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data used in Table S2 were from the 2015 China Household Panel Survey (CRHPS, 2015). Columns labeled Grains refer to the group of households that cultivated three major grain crops (rice, wheat and maize) only, and columns labeled Gross refer to the group of rural households that cultivated three major grain crops and/or three major cash crops (bean, peanut and rapeseed). The dependent variables in Table S3 are agricultural output per Ha and agricultural output per labor. Detailed explanations of the dependent and independent variables can be found in the section "Variables from the CRHPS 2015."

Summary Statistics for Tables S2 and S3

Variable	Obs	Mean	SD	Min	Max
Sowing area (grains, Ha, logarithm)	3370	-0.85	1.13	-5.01	5.30
Sowing area (gross, Ha, logarithm)	4030	-0.82	1.13	-5.01	8.29
Agricultural chemical use per sowing area (grains, Yuan/Ha, logarithm)	3220	8.11	0.96	1.28	13.47
Agricultural chemical use per sowing area (gross, Yuan/Ha, logarithm)	3865	8.08	0.98	0.95	13.47
Fertilizer use per sowing area (grains, Yuan/Ha, logarithm)	3169	7.89	0.94	1.28	13.47
Pesticide use per sowing area (grains, Yuan/Ha, logarithm)	2883	6.39	1.23	1.75	11.78
Contractual area (Ha, logarithm)	3918	-1.03	1.01	-5.30	5.30
Land parcels (categorical variable)	4053	0.64	0.48	0	1
Plant type (categorical variable)	4072	---	---	1	7
Land type (categorical variable)	4072	---	---	1	4
Land quality (categorical variable)	3960	---	---	1	5
Terrain (categorical variable)	4072	---	---	1	6
Output per Ha (logarithm)	3184	9.51	0.77	-0.58	13.26
Output per Labor (logarithm)	3195	7.95	1.37	-1.10	13.59

Table S4 The relation between agricultural intermediate (materials and services used for agricultural production such as chemicals, machinery, etc.) input and farm size based on the CHIP2002 data.

	Gross	Grains	Cash crops	Gross
Sowing area	-0.275***			
	(0.0196)			
Contractual area		-0.251***	-0.415***	-0.241***
		(0.0205)	(0.0330)	(0.0211)
Province	YES	YES	YES	YES
Constant	6.577***	7.250***	7.046***	7.189***
	(0.146)	(0.196)	(0.449)	(0.342)
N	7917	7023	6787	7214
F	68.10	52.69	114.89	61.85
adj. R^2	0.164	0.146	0.264	0.157

All data in this table are log-transformed to calculate the elasticity of dependent variables. Standard errors are reported in parentheses, with *** *p<0.001*. Data were from the Chinese Household Income Project 2002 (CHIP2002). Columns titled "Gross" refer to cultivation of both grain and cash crops. The dependent variable in Table S2 was intermediate agricultural cost per land area. Cultivated land area refers to a household's cultivated area in total, and contracting land area refers to a household's land area which was allocated under the HCRS. Province is provincial fixed effects. The coefficients were estimated by OLS.

Summary Statistics for Table S4

Variable	Obs	Mean	SD	Min	Max
Sowing area	8040	-1.11	0.84	-5.23	2.48
Contractual area	7318	1.60	0.83	-3.00	5.19
Intermediate input per sowing area (gross)	7917	7.77	0.92	0.92	12.36
Intermediate input per contractual area (grain)	7023	7.19	0.94	0.92	10.68
Intermediate input per contractual area (cash)	6787	6.47	1.49	0.36	12.36
Intermediate input per contractual area (gross)	7214	7.78	0.90	0.92	12.36

Table S5 Two-stage least squares (2SLS) estimation by instrumenting rural households' farm size (sowing area) with the contractual area, based on data from the CRHPS 2015

	Agricultural Chemical uses	
	Grains (IV)	Gross (IV)
Sowing area	-0.124***	-0.088**
	(0.024)	(0.023)
Land Parcels	YES	YES
Plant type	YES	YES
Land type	YES	YES
Land quality	YES	YES
Terrain	YES	YES
N	3068	3681
Wald chi2 stat	361.88	263.36
adj. R^2	0.163	0.111
Regression summary statistics for the first stage		
Contractual area	0.778***	0.751***
	(0.014)	(0.013)
Robust F stat for the first stage	2891.79	3234.04
Shea's Adj. Partial R-sq.	0.4827	0.4655

We used the contractual area as an instrumental variable to extract variations in farm size (sowing area) and fertilizer use intensity that are not correlated with farmer's knowledge and skills. The contractual size was largely determined before the early 2000s on an egalitarian basis. However, because the transfer of land use rights in China has been limited by various factors in the last decade, households' contractual area is strongly and positively correlated with their current farm size (sowing area).

Table S6 Sensitivity analysis of fertilizer use per area responding to socioeconomic factors

	Ln Farm size	PGDP	PGDP^2	Ln Price ratio	Beans%	Vegetable%
Changes of Ln fertilizer use	-1.84	+0.66	-0.45	-0.05	-0.50	+0.23
Equaling to how many standard deviation of fertilizer use changes	-1.30	+0.48	-0.32	-0.04	-0.36	+0.16

The results in the first row represent the changes in fertilizer use per area if each of the standardized socioeconomic factors changes by one standard deviation. The results in the second row represent a normalized changes within one standard deviation. Data used in this table are the same to those in Table 1. PGDP is per capita gross domestic product. Farm size represents the average farm size owned by rural households. Price ratio refers to the fertilizer to crop price ratio. Beans% represents the proportion of cultivated area of leguminous crops in total cultivated crop area, and Vegetable% represents the proportion of cultivated area of vegetables in the total cultivated crop area.

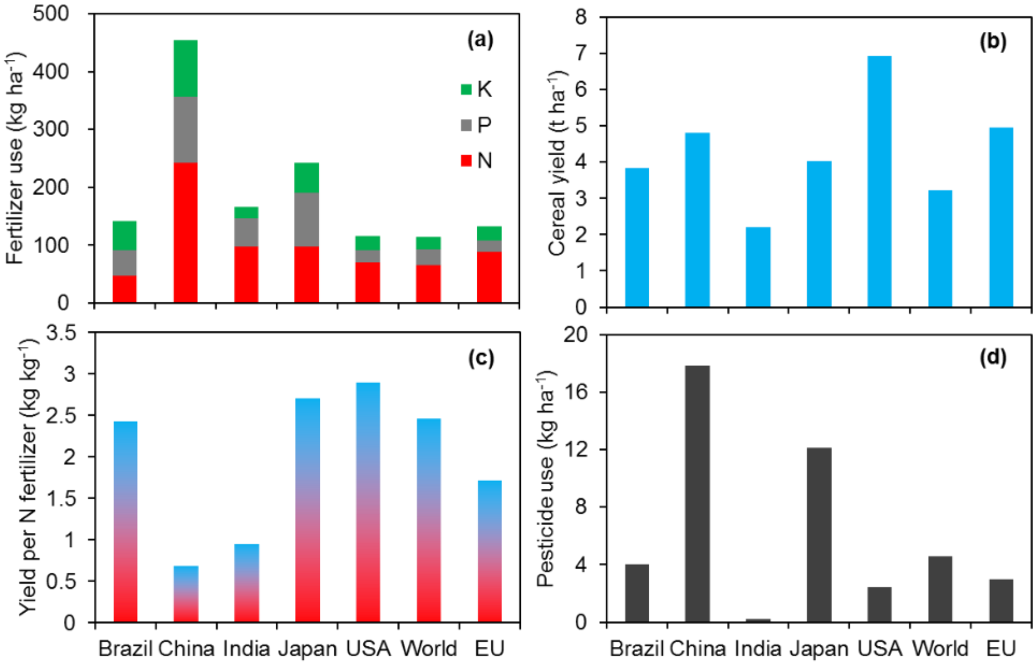

Figure S1. International comparison of agricultural chemical uses and their efficiency on crop yield. (a) Fertilizer use per area; (b) Cereal yield using rice milled equivalent; (c) Yield per N fertilizer use (considering wheat, maize and rice); (d) Pesticide use per area.

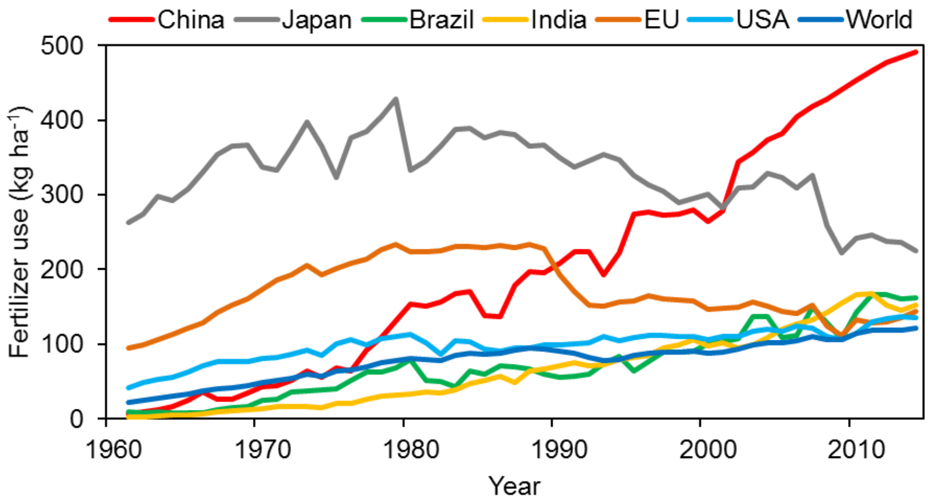

Figure S2. Changes in fertilizer use per area for major countries since 1960. Fertilizer use refers to total nutrients including N, P_20_5 and K_2O. Data used in Figure S3 were from FAO's food and agriculture database, available at www.fao.org. Time series data show that the overuse of fertilizer has only occurred in recent years in China. In the 1960s and 1970s, fertilizer use per area in China was much lower than many other countries. Fertilizer use per area increased rapidly after the late 1970s, which coincided with a period of rapid economic growth. Economic growth in many developed countries eventually led to a reduction in fertilizer use per area to reasonable levels. However, the use of fertilizer in China has continued to skyrocket as its economy grows, suggesting that there are factors preventing reduced fertilizer use in developed countries from being effective in China.

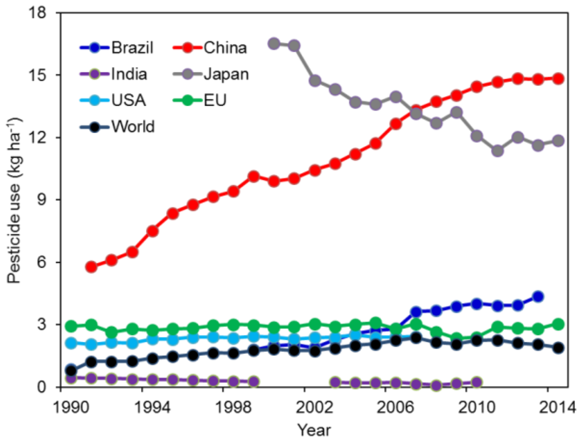

Figure S3. Country level pesticide use per arable land. Data used in this figure were from the FAO's food and agriculture database, available at www.fao.org

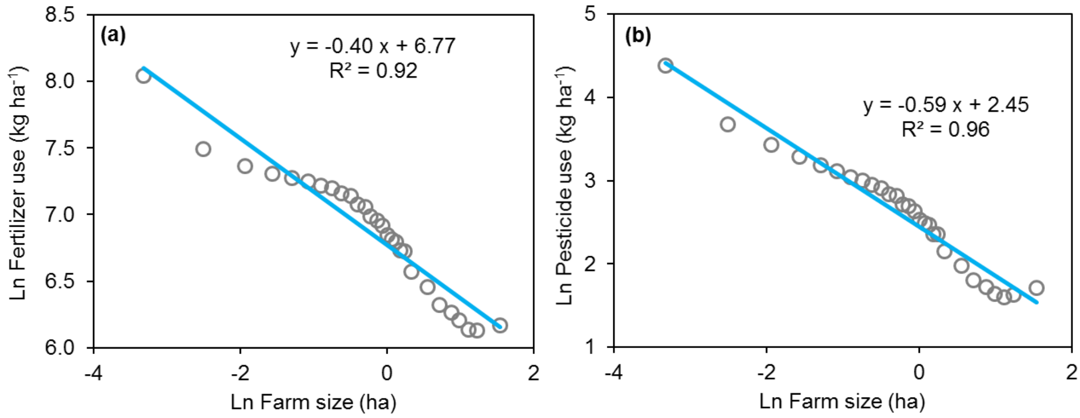

Figure S4. Fertilizer and pesticide used with farm size. Data used in Figure S4 were from the China's Second Agricultural Census (CSAC, 2009), which included all the smallholder farms (over 200 million households) and collective farms (over 395,000 collectives).

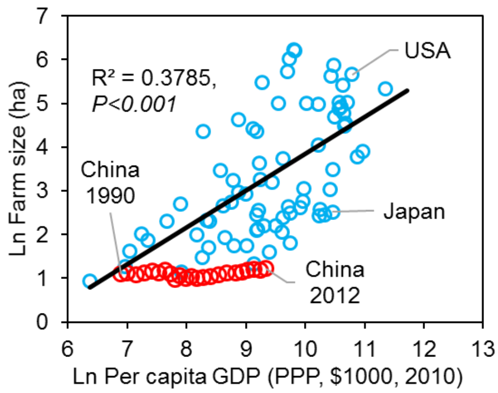

Figure S5. Farm size changes with per capita gross domestic product (PGDP) in different countries. Data sources were FAO and World Bank Open Data.

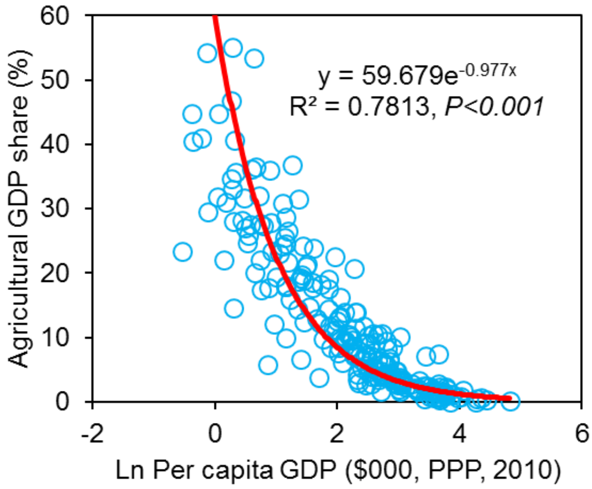

Figure S6. Decline of agricultural GDP share with the growth of per capita GDP. The data were from the World Bank Open Data, consisting of 193 countries in 2010. Purchasing-power-parity-based GDP per capita is used in the figure.

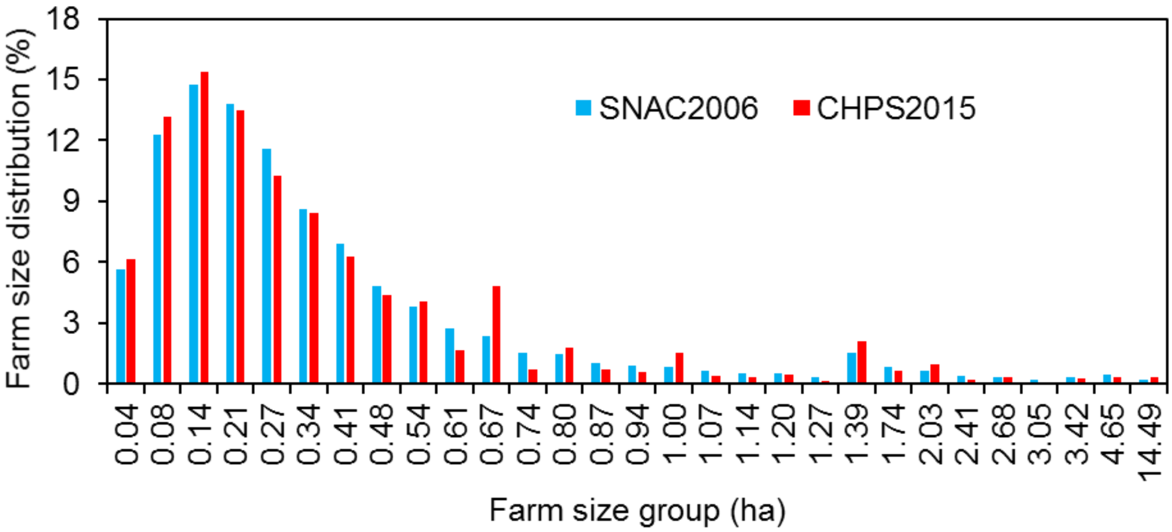

Figure S7. Comparison of farm size distribution in China between two independent surveys. The SNAC2006 refers to the Second National Agricultural Census in China in the year 2006, and the CRHPS 2015 refers to the China Household Panel Survey in the year 2015. 1 mu = 0.0667 hectares. The maximum group (14.49 ha) of CRHPS 2015 includes 71 households, and the distribution of farm size in Fig. S8 is adjusted for sampling weights.

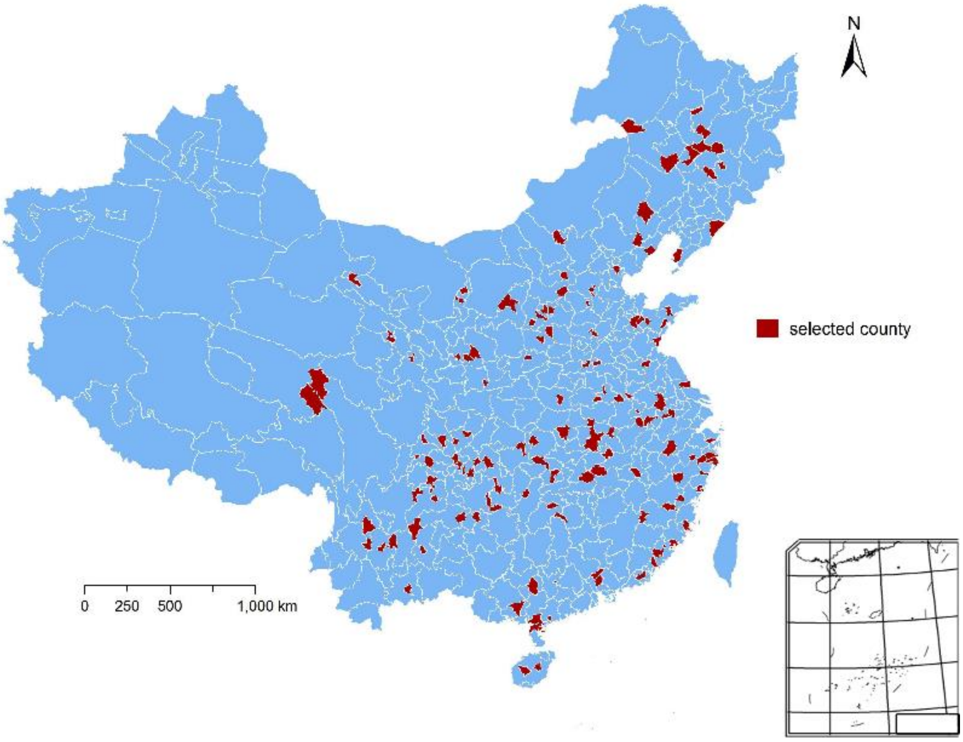

Figure S8. Locations of the selected counties in the 2015 China Household Panel Survey

References:

1. Lowder SK, Skoet J, Raney T (2016) The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development* 87:16-29.
2. Zhang X, Davidson EA, Mauzerall DL, Searchinger TD, Dumas P, Shen Y (2015) Managing nitrogen for sustainable development. *Nature* 528:51-59.
3. Gan L, Yin Z, Jia N, Xu S, Ma S, Zheng L (2014) *Data you need to know about China, Research Report of China Household Finance Survey · 2012* (Springer-Verlag, Berlin Heidelberg).
4. Eichenberger P, Hulliger B, Potterat J (2011) Two Measures for Sample Size Determination. *Survey Research Methods* 5:27-37.
5. Chen X, Cui Z, Fan M, Vitousek P, Zhao M, Ma W, Wang Z, Zhang W, Yan X, Yang J, *et al.* (2014) Producing more grain with lower environmental costs. *Nature* 514:486-489.
6. Cui Z, Zhang H, Chen X, Zhang C, Ma W, Huang C, Zhang W, Mi G, Miao Y, Li X, *et al.* (2018) Pursuing sustainable productivity with millions of smallholder farmers. *Nature* 555:363-366.
7. Gu B, Ju X, Chang J, Ge Y, Vitousek PM (2015) Integrated reactive nitrogen budgets and future trends in China. *Proc Natl Acad Sci USA* 112:8792-8797.
8. Sicular T, Ximing Y, Gustafsson B, Shi L (2007) The urban – rural income gap and inequality in China. *Review of Income and Wealth* 53:93-126.