Working paper

International Growth Centre

Arsenic Contamination, Nutrition and Economic Growth in Bangaldesh

Mark M. Pitt
Mark R. Rosenzweig
Nazmul Hassan

August 2011

When citing this paper, please use the title and the following reference number:
E-3026-BGD-1

DIRECTED BY LSE UNIVERSITY OF OXFORD

FUNDED BY UKaid from the British people

Arsenic Contamination, Nutrition and Economic Growth in Bangladesh

Mark M. Pitt
Brown University

Mark R. Rosenzweig
Yale University

Nazmul Hassan
University of Dhaka

August 2011

The research reported in this paper was supported by grants from the National Institutes of Health, ROlDK072413, the International Growth Center, RA-2009-ll-028, and from the MacMillan Center, Yale University. We are grateful to Katherine Kelley and Dave Murray for their excellent lab work.

In the 1970s and 1980s, the government of Bangladesh, with the support and financing of the United Nations Children's Fund, promoted the digging of tube wells to provide clean drinking water and reduce the incidence of diarrheal disease. Prior to this, drinking water came from surface water sources, which were identified as a principal source of diarrheal disease. In the late 1990s, evidence indicated that groundwater, unlike surface water and by then the main source of drinking water, irrigation and water for cooking in Bangladesh, was contaminated by naturally-occurring arsenic in 59 of the country's 64 districts. The contamination of groundwater by arsenic in Bangladesh is the largest poisoning of a population in history (Smith *et al.*, 2000). As 95% of the country's population relies on well water, an estimated 57 million people have been chronically exposed to drinking water with arsenic levels exceeding the WHO standard of 10 μg/L, and 35 million people were exposed to arsenic levels above the country's government standard of 50 μg/L (British Geological Survey, 1999).

In this paper we provide new estimates of (i) the effects of the consumption of foods grown and cooked in arsenic-contaminated water on individual arsenic concentrations and (ii) how the ingestion and retention of inorganic arsenic causally affects nutritional status, physical and cognitive capabilities and incomes at the individual level. The estimates are based on arsenic biomarkers obtained from a sample of members of rural households in Bangladesh who are participants in a long-term panel survey following respondents and their coresident household members over a period of 26 years. In the last round of the survey each respondent in the survey aged eight years and above provided clippings from all ten toenails. Toenail clippings are the preferred biomarkers for measurements of arsenic ingestion in environments in which exposure lasts more than a few months and where a significant share of arsenic exposure occurs through food consumption. Our new estimates of the causal impacts of arsenic exploit the known genetic linkages among kin in their ability to methylate arsenic, thus reducing its toxic effects, and data identifying family linkages among respondents living in separate environments (villages).

The link between arsenic exposure and an individual's health and productivity has two components: ingestion and retention. Arsenic is ingested by drinking and by consuming foods grown and cooked in arsenic-contaminated waters. Estimates of the relative importance of the two sources, based on the distribution of arsenic among wells in Bangladesh and the non-linear relationship between arsenic ingestion and exposure levels, suggest that dietary sources of arsenic are the most important route of arsenic exposure for two-thirds of the Bangladesh population (Kile *et al.* 2007). The toxicity and retention of arsenic that is ingested, however, is also affected by the nutrients contained in the food consumed. Specific foods differ in the degree to which they retain arsenic used in their cultivation and cooking and differ as well in the extent to which, via their nutritional composition, they enable the metabolism

(detoxification) and subsequent excretion of arsenic, primarily through a process of methylation. For example, Zablotska *et al.* (2008) and Heck *et al.* (2007), based on randomized field experiments, highlight the protective effects of folate-related nutrients and B-vitamins in samples from rural Bangladesh. Leafy vegetables are known to contain high concentration of folic acid. However, Hossain (2006), based on extensive tests in Bangladesh, reports that among the food products tested, leafy vegetables have the highest concentration of arsenic. The net effect of the consumption of different foods on arsenic stored in the body is thus not known.

A large number of studies have related arsenic exposure in the environment, measured by well contamination, and diet to biomarkers for arsenic in humans (urine, blood, hair, and toenails) and to specific health outcomes. However, these studies have a number of deficiencies. First, neither proximity to arsenic sources (wells) nor the ingestion of arsenic can be considered exogenous to the outcomes studied, especially as arsenic ingestion is related to the nutrient composition of foods, which reflect preferences and income constraints. Moreover, studies of actual food intakes on arsenic ingestion and on health do not take into account the endogeneity of diets. An individual's retained arsenic in a country such as Bangladesh, where millions of wells have been tested and identified by their levels of arsenic contamination and where diet is both a source of arsenic and an aid to its metabolism, is clearly not an exogenous variable. Second, many studies employ only women as subjects. Evidence suggests that women have better methylation efficiency than men, although the women are also less likely to consume the foods that would aid in methylation, consistent with laboratory findings that suggest a significant protective effect of estrogen on arsenic methylation (Lindberg *et al.* 2007, 2008). These studies may therefore underestimate the effects of arsenic exposure for the general population.

Third, studies that carry out randomized nutrient supplementation that can reduce arsenic's burden by increasing the rate at which it is metabolized (methylated) do not identify the mechanisms by which such arsenic-reducing interventions improve health, productivity and well-being. Such supplements improve health and productivity by more than just reducing the toxic concentration of arsenic and its metabolites in body tissue. Knowledge of both how arsenic retention directly and causally affects health and productivity outcomes and how diet affects arsenic retention is obviously critical to any cost-benefit analyses of governmental programs aimed at improving health and productivity. Fourth, the effects of such ameliorative interventions may differ by the level of *As* concentrations, but most existing studies have been based on small samples that preclude the identification of nonlinear effects. Knowledge of which population groups would most benefit from such interventions is obviously useful in setting policies that seek to ameliorate the effects of arsenic contamination in the environment. Identification of

how arsenic ingestion causally affects health and productivity thus must take into account all of the fundamental determinants of its toxicity, which include behavioral, environmental and genetic components as well as the distribution of arsenic concentrations in the population.

In section 2 of the paper we set out our methodology for both estimating the effects of diet on arsenic retention and identifying the effects of arsenic retention in the body on health and productivity measures. Section 3 describes the data and our measure of arsenic retention from the sample of toenails and provides descriptive statistics on the levels and distributions of arsenic concentration in the sample. The data indicate very high levels of arsenic concentrations, exceeding average levels measured in US respondents by almost 20 to one, with concentrations exhibiting wide individual variation but spread almost uniformly across education and landholding groups. In section 4 we provide estimates of the effects of the individual consumption of seven food groups on arsenic concentrations based on individual-specific food consumption information and the toenail-based arsenic measures. These show that diet matters for arsenic contamination, with the consumption of fruits, vegetables and leafy vegetables all having significant positive effects on measured arsenic concentrations. We show that increases in wealth have small, and diminishing with income levels, effects on the consumption of all three food groups, indicating that economic growth, in the absence of direct initiatives to reduce arsenic exposure, will have little impact on the toxic effects of arsenic.

Section 5 contains the estimates of the effects of individual-specific arsenic concentrations on measures of cognition, nutrition status, schooling and, for males, daily wage rates. We find that increased arsenic retention (i) significantly reduces weight, body mass, and school participation for males and females equally, (ii) significantly reduces cognitive ability, as measured by performance on the Raven's CPM test, and strength for men but has little effect on either capability measure for women, and (iii) significantly reduces male wage rates. Our estimates also indicate that reductions in arsenic concentrations have nonlinear effects, with equal reductions of arsenic retention having greater effects at relatively low, but still elevated, levels of arsenic concentrations than at the highest levels in the Bangladesh population. This means that interventions that marginally improve arsenic methylation such as supplementation with folate will be more effective in significantly reducing the productivity effects of arsenic in rural Bangladesh if targeted to groups with lower not higher concentrations of arsenic. Finally, we look at wealth effects on food consumption. We find that the composition of the diet, and the consumption of vegetables and fruits that our estimates indicate increase *As* concentrations, are insensitive to large changes in wealth and as a consequence *As* is also unaffected by wealth. Given that income growth alone has little effect on population arsenic concentrations, policies that eliminate the

source of arsenic exposure maybe most cost-effective, particularly in areas where such concentrations are high and thus where ameliorative interventions have the least efficacy.

1. Method of analysis

The main problem of interest in our study is to identify how arsenic concentrations that result from the ingestion of arsenic affect health and productivity and to identify how changes in food consumption, by affecting arsenic ingestion and excretion, affect *As* concentrations. More formally, we wish to identify the parameter δ in the policy function, or conditional demand equation, for an individual i in household j:

$$\mathbf{(1)} \qquad h_{ijk} = \delta A_{\mathbf{ijk}} + \boldsymbol{Z}_{\mathbf{kij}}\boldsymbol{\beta}_z + u_{ijk} + u_j + \varepsilon_{ijk},$$

where $A_{\mathbf{ijk}}$ is the individual's stored concentration of arsenic , h_{ijk} is some human-capital or productivity outcome such as cognitive ability, grip strength, school enrollment, labor supply or wage rate, $\boldsymbol{Z}_{\mathbf{k}ij}$ is a vector of observed exogenous attributes of the individual and household, u_{ijk} an individual-specific error, u_j is a household fixed effect (reflecting, for example, the local the health environment), and ε_{ijk} an iid error. The k index identifies the individual's relationship with a family member (lineage). δ is the effect of changing an individual's arsenic "toxic burden," measured by her arsenic concentration, on productivity and health. The problem for estimation is that A_{ijk} may be correlated with the unmeasured household and individual-specific variables associated with h_{ijk} so that least squares estimation of (1) would not provide a consistent estimate of δ. For example, a household or individual particularly concerned about health may avoid contaminated wells (which are identified in many villages), thus spending more time than others in water collection and thus perhaps less in earning activities. Such a person may also, unaware of the different levels of arsenic contamination of food groups, consume more contaminated foods in order to improve health (e.g., more green vegetables and fruits) or commit other resources to activities or inputs that affect human capital via health.

Randomized interventions at the village level, such as the construction of a deep well, cannot be used as an instrument for individual arsenic stock because the new well will directly alter the allocation of time of all family members, for example, and thus affect labor market and other outcomes directly for any individual (e.g., children spend more time fetching water and less time doing homework). Even randomized individual-specific interventions that reduce arsenic cannot usually be used as instruments to identify δ. For example, the randomized distribution of nutrition supplements across individuals, which have been shown to increase arsenic metabolism and excretion (as in Gamble *et al.* 2007), would have direct effects on health and productivity in addition to any effects via the stock of arsenic. Such an intervention thus would not identify the arsenic-health or productivity mechanisms. What is required is exogenous variation in a variable that directly affects arsenic in the body, given a person's exposure to arsenic in the environment, and has no other direct effects on the outcomes of interest.

Our strategy for the identification of δ exploits genetic linkages among family members that influence an individual's ability to metabolize arsenic. Recent evidence indicates that genetic variations are a source of exogenous variation in the arsenic stock of the body, and hence the toxicity of any ingested quantity of arsenic, and form the basis for an IV estimation strategy. This evidence suggests that, apart from the nutritional determinants of arsenic concentrations in humans, genetic polymorphisms are a major cause of the substantial inter-individual variation in arsenic metabolism (methylation) within the same exposure area (Vahter, 2000). For example, Kile *et al.* (1995) finds that individuals in Bangladesh possessing GSTT1-null genotypes had significantly more arsenic in their toenails in contrast to GSTT1 wild-type individuals, after controlling for drinking water contamination. Also working in Bangladesh, McCarty *et al.* (2007) find that the *GSTT1* wildtype modifies the risk of skin lesions among arsenic-exposed individuals. Ahsan *et al.* (2007) find that variations in the MTHFR and GSTO1 genes modulate health effects of arsenic exposure, noting that the consistent relationships of the risk of skin lesions with genotypes and diplotypes in the MTHFR and GSTO1 genes suggest the importance of genetic susceptibilities in their Bangladeshi study population.

The genetic origins of arsenic metabolism suggest that the ability to methylate is correlated among family members. For example, in Chung *et al.* (2002) families from Chile were selected based on their long-term exposure to very high levels of arsenic in drinking water (735–762 μg/L). Each family consisted of a father, a mother, and two children. Urinary arsenic and its methylated metabolites for each participant was measured. The intra-class correlation coefficients showed that 13–52% of the variations in the methylation patterns were from being a member of a specific family. Family correlations were calculated for father–mother, parent–child, and sibling–sibling pairs. Methylation patterns correlated strongly between siblings compared to lower correlations in father–mother pairs, after adjustment for total urinary arsenic, age, and sex. Obviously these correlations could arise due to similarity in diets and to common environmental factors. However, the intra-family correlations were not notably altered when adjustments were made for specific blood micro-nutrients (methionine, homocysteine, folate, vitamin B6, selenium, and vitamin B12) potentially related to methylation. However, a superior methodology for identifying genetic correlations would (i) remove from the measures of arsenic concentrations that part due to diet variation and (ii) examine such correlations among family members not residing in the same household or even the same environment. Our data enables us to do this.

We use as an instrument for an individual's stock of arsenic the individual's genetic ability to metabolize (methylate) arsenic based on a non co-resident family member's ability to do so net of common factors in the environment. This requires that we identify among genetically-linked, non-coresident individuals that component of A_{ijk} that is unexplained by exposure to either environmental arsenic or by endogenously-determined individual-specific nutritional intakes.

To carry out our method, we thus need to first estimate the production function for arsenic

concentrations, based on the measures taken from the toenail clippings:

$$\textbf{(2)} \qquad A_{ijk} = \boldsymbol{N}_{ijk}\boldsymbol{\alpha} + \lambda_m\mu_j + \mu_k + \mu_{ijk} + e_{ijk},$$

where $\boldsymbol{N}_{ijk}$ is a vector of person-specific biological inputs, such as food consumed, and $\boldsymbol{\alpha}$ is a vector of coefficients. Equation (2) contains four sources of unobserved heterogeneity: The first is the genetic component that is shared among a lineage or kin group μ_j according to the degree of shared genes, as reflected in the factor-loading for λ_m, where the subscript m indicates genetic distance – in order, siblings, parent or child, uncles or aunts, other relatives. There is also a household component that represents the unobserved exposure to environmental arsenic from, say, household drinking and cooking water, μ_k, an individual-specific idiosyncratic fixed effect μ_{ijk}, and an iid error term e_{ijk}. All of the μ components capture in part the documented inter-individual genetic variation in the ability to metabolize inorganic arsenic. The more efficient an individual is at metabolizing arsenic, the less the body's toxic burden and the more arsenic is quickly eliminated in urine and not deposited in tissue. The μ_j captures that part of the genetic variation that is correlated within lineages.

There are two challenges to identification of the genetic component of (2). The first is to obtain consistent estimates of the $\boldsymbol{\alpha}$, the effects of variation in diet on the arsenic stock. Food intakes are clearly endogenous, correlated with unobservables such as preferences for health or leisure or work type that are also impounded in the μ_k and μ_{ijk}. To eliminate the household component of the unobservable that may be correlated with food intakes, we use a family fixed effect estimator. However, individual food intakes will still be correlated with individual error components, as we and others have shown (e.g., the inherently strong receive more calories; Pitt *et al.* 1990). We therefore use as instruments for food and nutrient intakes the prices of foods, at the village level, interacted with the age and gender of the individual. Households reside in over 350 villages, so there is ample cross-sectional variation in food prices. This combination of IV and FE has been used in Pitt *et al.* (2000) to identify with reasonable precision the effects of the individual nutrient intakes on health outcomes such as diarrheal disease and respiratory ailments. Estimation of the nutrient effects on individual-specific arsenic concentrations thus exploits the real variation in relative food prices across the large number of villages in our sample, the existence of extended families in Bangladesh, and the information our data contain on individual food intakes.

The estimates of the $\boldsymbol{\alpha}$ are of direct interest. First, we can quantify to what extent the individual consumption of fruits and leafy vegetables, for example, which the literature suggests are a source of arsenic ingestion and also mediate the effects of arsenic ingestion via arsenic metabolism and excretion, on net affect arsenic concentrations in the body, given exposure to arsenic via water sources (impounded in the household fixed effect). For these estimates to be credible based on the short-period information on food intakes, it is important that the measure of individual arsenic concentrations, based on the toenail assays, reflects relatively recent, and not lifetime, accumulation, so that it is not necessary to have the complete life history of food intakes. The toenail-based measure mainly reflects arsenic ingestion in recent months, not years (Kile *et al.*,

2005).[1]

The "own" residuals from (2) cannot be used as instruments for an individual's arsenic concentration in the policy response function (1) because they contain the household fixed effect, which is likely to be correlated with the household effect u_k in the health outcome and productivity equation (individuals near a contaminated well may be more likely to work longer). The second challenge is thus to remove the household component from the compound error term, and to identify the λ_m factor-loadings. We make use of the fact that for almost every household in our original 1982 sample a relative had left the household between 1982 and the second round of the survey in 2002. And, between 2002 and the latest round, some proportion of households in the larger new sample of households added in 2002 also had a family member depart. Household division is mostly due to marriage - between 1982 and 2002, for example, 85% of girls age 2-14 in 1982 had left their original household and village and 10% of the boys left the village. Because as part of the survey design we include both the originally-sampled households and any household containing any individual from the original households no matter where the location in Bangladesh, we have as much information on family members who remained together as we do for those who are no longer co-resident.

The residuals from (2) for genetically-linked but non co-resident family members can be used as instruments for A_{ijk} in (1) to identify δ. That is, after estimation of (2), it is straightforward to compute $E(\mu_j|A_{ijk} - \boldsymbol{N}_{ijk}\boldsymbol{\alpha}$ for all members of lineage j, $\lambda_m, \sigma_{\mu k}, \sigma_{\mu j}, \sigma_{\mu ijk}+\sigma_{\varepsilon ijk})$. However, if there is measurement error in $\boldsymbol{N}_{ijk}$, then in lineages with small numbers of observations, this expectation will also be measured with error that may not be of the classical variety in subsequent estimation. If this measurement error is uncorrelated across member in a lineage, then it will be useful to estimate the expectation $E(\mu_j|TN_{ijk} - \boldsymbol{N}_{ijk}\boldsymbol{\alpha}$ for all members of lineage j except person ijk, $\lambda_m, \sigma_{\mu k}, \sigma_{\mu j}, \sigma_{\mu ijk}+\sigma_{\varepsilon ijk})$.

Identification requires that the household components of the error terms across households not located in the same village are uncorrelated. The short-run nature of the *As* concentration measure is again important - the fact that the separated family members once shared a common arsenic source should not be reflected in the toenail-based measure of *As* as long as the family members divided at least a year prior to the survey. An additional requirement for identification is that $E(\mu_j, u_{ijk}) = 0$. A necessary condition for this zero covariance is that the genetic polymorphisms that regulate the efficiency of arsenic metabolism are unrelated to those that affect h_{ijk}. For example, the cognitive ability endowment not be correlated with the arsenic metabolism

[1]Kile *et al. (*2005) collected toenail and drinking water samples from forty-eight families (n=223) every three months over a two year period. The team created a weighted, lagged exposure variable including drinking water arsenic concentrations measured three, six and nine months before toenail collection. Their findings suggest that the drinking water concentrations at three, six and nine months contributed 69%, 14%, and 17%, respectively, to the arsenic detected in the toenail sample.

endowment. Finally, our identification strategy requires that the genetic linkages λ_m are non-trivial.[2]

2. Data and Measurement of Arsenic Concentrations

a. Survey design and sample size

The data set consists of information on 13,258 individuals residing in 2,480 households from the 2007-8 round of a panel survey that started in 1981-2 in 15 villages in rural Bangladesh. The first survey was obtained using a sample that was meant to be representative of the rural population of Bangladesh (Ahmad and Hassan, 1982). The second round of the survey in 2002-3 included all of the individuals in the 1981-2 households residing in 14 of the 15 original villages, wherever they were located in Bangladesh, plus all of the individuals residing in any new households formed since the original round. Less than 3% of non-deceased individuals from the first round were not found and included in the survey second round. In addition, at the time of the second round of the survey a new random sample of households in the original 14 villages were surveyed, and in the latest 2007-8 round all of the individuals from the 2002-3 round and any new household members were included as respondents. Attrition for this round was less than 8%.

Because of the panel survey design, which tracked all individuals who were ever sampled regardless of location kinship relationships with other sampled individuals but residing in different households and villages can be identified. The spatial separation of kin who once co-resided results from migration. As is typical in South Asia, female household members leave their origin households at marriage. In the 2002-3 round, for example, 85% of women who were aged less than 15 had left their village birthplace. 10% of the males in that original age group also were residing in a village differing from the one in which they were born 20 years after the initial survey. In the 2007-8 round of the data the number of villages represented in the sample had grown because of migration from 14 in 1981-82 to 612. The information on family links among spatially-separated respondents, as noted, is key to the identification of the causal effects of arsenic ingestion and retention, given the genetic component of arsenic methylation.

The survey data provide more detailed information at the individual level than do most large-scale surveys. The data include individual-specific information on food intakes, observed by investigators over a 24-hour period. This information is required to carry out our estimation strategy. There is also a wide variety of individual-level productivity and health measures, including information on anthropometrics, labor force participation and wage rates. In the last round of the data measures of physical capabilities and cognition were

[2]We can also identify the λ_m. It is easy to show that identification of the λ_m requires that the sample contain not only individuals belonging to the same lineage who are not co-resident but also contain individuals who are not in the same lineage who are co-resident (spouses). More formally, to identify the λ_m factor-loadings one can estimate (2) using a variant of a two-factor random effects model with a specified correlation structure (generalized estimating equation (GEE)), with instrumental variables. Note that the λ_m are estimable even if the data contain only the maternal or only the paternal lineage as long as the is no assortative marriage on arsenic metabolism genes and the λ_m do not depend on whether the lineage is maternal or paternal, that is, metabolism genes are not sex-linked. Even then, there is no problem in estimating separate sets of λ_m for maternal kin and paternal kin. We do not pursue this here, because our lab analyses have been confined to lineage groups.

also obtained for every respondent meeting minimum age requirements. Among these are pulmonary function (measured by spirometry), symptoms of respiratory and other illness (based on clinical exam), cognitive and intellectual ability (measured by an abridged version of the Raven's Colored Progressive Matrices exam, and grip and pinch strength.

b. Measurement of respondent arsenic concentrations

A key component of this study is the measurement of *As* concentrations of the respondents. We collected from each respondent aged eight years and above clippings from all ten toenails, providing a basis for obtaining a contemporaneous measure of arsenic ingestion for 7,356 individuals. To extract information on arsenic concentration from toenails, trace metal analysis using inductively-coupled plasma mass spectrometry (ICP-MS) must be used after preparation, including cleaning with solvents and digestion in acid.[3] Toenail clippings are the preferred biomarkers for measurements of arsenic ingestion in every case in which exposure lasts more than a few months, and is particularly preferred when a significant share of arsenic exposure occurs through food consumption (Kile *et al.* 2007). Because of the considerable expense of the laboratory analyses and preparation, we assayed only a subset of the toenails. In particular, we confined the analyses to pairs of households containing related kin residing in different villages that have at least two, same-sex respondents within each pair. Although it is useful to have this information for all respondents for which we have these biomarkers, our identification strategy, as noted, requires data on arsenic concentrations for closely-related relatives living apart. Because of budget limitations, the number of over-eight respondents for whom we have arsenic concentration (*As*) measures is 4,260. Due to the sampling strategy used for the laboratory analyses, the sample of respondents with the *As* information over-represents married women, due to their relatively greater mobility, and is somewhat older than the sample of respondents with unanalyzed toenails. Neither average schooling levels nor landholdings, however, differ between the two subsamples. In all of our empirical analyses using this information, we will separately analyze males and females and control for age (and landholdings).

As part of the assessment of the laboratory methods, a sample of 25 respondents (graduate students residing in Rhode Island) also provided toenail clippings. These samples provide not only a basis for assessing the capabilities of the laboratory methods used to measure low levels of *As* concentrations, but also a benchmark with which to compare the levels measured from the Bangladesh sample. The analyses indicate that the concentrations of arsenic in the Bangladesh respondents are quite high, vary considerably, but are spread across all economic groups. In the US graduate student sample, average *As* concentrations are 78.3 parts per billion (ppb), with a standard deviation of 46.6 . In the sample of Bangladesh respondents the average

[3]Appendix A provides a detailed description of the methods that were used in the ICP-MS Laboratory at the Graduate School of Oceanography at the University of Rhode Island.

concentration is 1,353 ppb, with a standard deviation of 1,894.[4] Figure 1 provides the frequency distribution of the arsenic concentrations measured in the two samples, which show the substantial contamination of the Bangladesh respondents - 90% of the Bangladesh sample have *As* concentrations greater than the highest value found in the US sample, and over a third have concentrations exceeding 1,000 ppb. Figures 2 and 3 show that arsenic contamination is not confined to the poor and least educated in Bangladesh. In Figure 2, which displays levels of *As* concentrations by owned landholdings (area), arsenic contamination is slightly more elevated among households with larger landholdings. Figure 3 shows that *As* concentration is uniform by educational level up through eight years of schooling, which represents over 80% of the sample, and is only slightly lower for respondents with more than eight years of schooling attainment.

c. Analysis sample

For the analysis of the effects of individual *As* concentrations on nutritional status, capabilities and earnings, we use a sample of respondents aged 15-29. We selected this sample because they were almost surely exposed uniformly to arsenic through drinking over the course of their entire lives, unlike older cohorts, and thus are more representative of future cohorts in Bangladesh in the absence of policies that eliminate arsenic contamination. In the 1981-82 round of the data, one-third of respondents drank water from sources other than tubewells; in the 2002-3 round, tubewells were the source of drinking water for 97.6% of respondents. Water used for cooking was obtained from tubewells for 72.2% of respondents. For this subsample, we have the *As* concentration measure for 52.3% of female and 44.5% of male respondents, the imbalance due to the oversampling of toenail samples for split family members and the higher mobility of women. Respondents in the subsmaple reside in 465 villages. Based on the kinship relationships, we constructed 583 lineage groups - respondents living in different villages who are either a sibling or a parent-child pair. Table 1 provides average *As* concentrations and food intakes by gender for the subsample. As can be seen, women have concentration levels that are higher, by 14.7%, than men, and also appear to have different diets - women consume 27% more green vegetables than do men and 12.6% more tubers.[5] In Table 2 we also see that women do less well on the Raven's test, and are less likely to attend school (in the age range 15-23). Women on average are also less strong and smaller than men, but that is true in all human populations.

3. Diet and Arsenic Concentrations

We first estimate equation (2), the relationship between food intakes and arsenic concentration, using the individual-level information on food consumption, divided into seven food groups - grains, pulses, green vegetables, other vegetables, tubers, fruits and meat, fish and dairy. As noted, the toenail-based *As*

[4]These sample differences in arsenic concentrations are not due to differences in the age of the respondents. For a comparably-aged subsample (22-30 years of age) the mean and sd statistics for the Bangladesh respondents are 1,365 and 1854, respectively.

[5]These diet composition differences are also in evidence in the full sample.

concentration measure represents stored arsenic in the body reflecting arsenic ingestion over the past three months, while the food intakes are measured in a 24-hour period. Both the outcome and input variables thus are short-term, but the food intake variables measure with error the food consumed over the period relevant to the concentration measure. For this reason, and because unobservables that affect diet may also affect the choice of foods, we use instrumental variables estimation method, as in Pitt *et al.* (1990, forthcoming), employing the set of village-level food prices as instruments.

Table 3 reports OLS and IV estimates of the diet-arsenic production function. All food variables are expressed in logs, as is the concentration of arsenic. While the signs of the OLS and IV coefficients are identical, the OLS coefficients for all but meat consumption are biased towards zero. The Anderson test for underidentification indicates strong rejection of the underidentification null. The hypothesis that the set of OLS and IV coefficients are identical is also rejected ($\chi^2(9)=80.3$). The estimates of the gender effect, using either estimation method, indicate that women, net of dietary intakes, retain 10% less arsenic in their bodies than do men. This result is consistent with the medical literature indicating that women methylate more than do men as a consequence in part of the protective effect of estrogen (Lindberg *et al.* 2007). That in our sample, on average women have more *As* concentrations than men thus appears to be because in part women consume different diets than do men.

The IV estimates indicate that the consumption of three food groups significantly increase the retention of arsenic in the body in the contaminated water environment of Bangladesh - green and other vegetables and fruits. Recall that green vegetables, based on evidence obtained from the randomized distribution of folate supplements, given arsenic contamination, lower arsenic storage in the body because of increased methylation. But green and other vegetables also increase arsenic ingestion, as they retain arsenic-contaminated cooking water and arsenic contained in contaminated water used for irrigation. The positive coefficient on green vegetable thus reflects the net effect of the negative arsenic-mitigating effects of folate and the retention by vegetables of contaminated water. The effects of green vegetable intake, however, is less than that of green vegetables, consistent with the folate concentrations of the latter group, but both worsen *As* contamination in the body. The IV point estimates indicate that increasing the amounts of green and other vegetables consumed by one standard deviation would increase arsenic contamination by 14.9 and 10 percent, respectively. The other coefficient that is measured with precision, that for fruits, indicates that increasing the consumption of this food group by one standard deviation at the sample mean would increase arsenic concentration by 29%.

4. The Effects of *As* Concentrations on Cognitive and Physical Capabilities, Nutritional Status and Schooling

a. Linear and log-linear IV estimates

The correlation between a respondent's *As* concentration and her productivity and health does not provide an estimate of how a reduction in *As* would affect such outcomes because it reflects both food consumption choices and environmental factors that may jointly affect *As* ingestion and outcomes. The

residuals from the estimates of the *As* production function eliminate diet choice as a factor affecting *As* storage. They contain both the individual-specific ability to methylate arsenic as well as those local environmental factors that affect arsenic ingestion not due to food consumption. The residual measure of *As* concentration thus cannot be used directly to estimate the effect of stored *As* on outcomes. We use as an instrument for own *As* the average of the *As* residuals of all same-sex siblings and parents of the respondent (the respondent's lineage) living in another village. Because the gene for methylate ability is linked across kin, the lineage-based instrument should be correlated with the respondent's own storage of arsenic, but the not with local environmental factors affecting the respondent's arsenic ingestion. Even though all of the identifiable kin were once living together in the same environment, because the toenail-based measure of *As* reflects only the past three months of ingestion and retention, the influence of the past common environmental factors is absent.

Table 4 provides IV estimates of the effects of arsenic concentrations on direct measures of the respondent's cognitive and physical capabilities, specifically on performance on the abridged Raven's CPM (number of questions answered correctly) and on the log of pinch and grip strength test results (kg of pressure) using standard dynometers, for the full sample of respondent's aged 15-29 and by gender.[6] Also included in the specification are the respondent's age and its square, the value of own household landholdings and the number and average ages of co-resident adult males and females. Table A in the Appendix provides the first-stage estimates of the relationship between the non-coresident lineage residual *As* measure and own *As* concentration. The lineage measure is strongly correlated with own *As* despite the spatial separation, consistent with the medical literature. The power of the lineage variable is seen in the Anderson LM test statistics in Table 4, which strongly reject the hypothesis that the estimates are underidentified.[7]

The results in Table 4 indicate that the ingestion and retention of arsenic in the body significantly reduces cognitive ability and strength for males but not for females. The point estimates indicate that a one standard deviation increase in *As* concentration reduces the Raven's performance for males by a third of a question (6%). To put that estimate in perspective, we can compare the effect to that resulting from a change in the value of a household's land using the precisely estimated, and positive, land value coefficient in the Raven's test specification. The land value point estimate suggests that it would take a 1,000,000 *taka* (US$14,000) increase in land value, more than double the mean land value in the sample, to increase test performance by the same amount as reducing arsenic concentrations by one standard deviation.

The estimates for performance with respect to strength indicate that an increase in *As* concentrations by

[6]Performances on the Raven's test have a symmetric distribution, while the distribution of the strength test outcomes are strongly skewed to the right. Results using a linear specification rather than the more appropriate log-linear specification for the pinch and grip tests are similar to those reported.

[7]Tables B-D in the Appendix provide the IV and OLS estimates of the effects of *As* on all outcome measures. In all cases but two - weight for males and schooling attainment for females - the OLS estimates substantially understate the impact of *As* concentrations on outcome measures relative to their IV counterparts.

one standard deviation reduces male pinch and grip performance by 3.3% and 2.1%, respectively. The corresponding estimates for women, all of which are not statistically significantly different from zero, are one tenth of a question increase in correct responses (3%) and less than an 0.1% increase for either strength measure. The lack of any effect of *As* concentrations on strength for women is consistent with the medical literature, which indicates that nutritional intake significantly affects the strength (brawn) of men but not women, as is verified using a similar sample from the survey data in Pitt *et al.* (forthcoming).

As concentrations also appear to reduce nutritional status. In contrast to the results for cognition and strength, increased concentrations of *As* affect measures of nutritional status equally for males and females. This again is consistent with the gender findings on strength - equal reductions in nutritional status for men and women results only in reductions in strength for men. The point estimates here indicate that a one standard deviation increase in arsenic storage only reduces height by .62 centimeters, less than 0.4%, reflecting the short-term nature of the *As* measure being used. However, weight is reduced by almost one kilogram (2%) and weight-for-height by 1.6%. The point estimates for women are higher than those for men for weight and weight-for-height, and their mean values are also lower. As a consequence the impact of arsenic concentrations on these anthropometric measures is slightly higher in percentage terms for women: a one standard deviation in *A*s reduces weight and weight-for-height for women by 3% and 2%, respectively.

As concentrations also appear to affect school attendance, but not school attainment, reflecting again the short-term nature of the biomarker. For the estimates of the effects on school attendance we use a slightly younger sample - ages 15-23 - as no one in the sample over age 23 is attending school. For this group, although the negative effect on school attendance is larger for males, the effect is not statistically significantly different by gender. The point estimate for the combined sample indicates that a one standard deviation increase in the concentration of arsenic reduces the probability of being in school by six percentage points, or by over 75%, as only eight percent of respondents in this age group are in school.

5. IV Estimates of *As* Concentrations on Male Wage Rates

The findings that increased concentrations of arsenic retained in the body reduce strength, contemporaneous nutritional status and schooling suggest that respondents with arsenic-contaminated bodies will suffer an earnings loss, as all three are key determinants of earnings in rural Bangladesh (Pitt *et al.*, forthcoming). Although few women work for wages in rural Bangladesh, 72% of the sample of males in the age group 15-29 worked sometime during the year for wage rates. The data provide wage earnings and days worked by month, and we construct a daily wage measure that is the mean of the per-month daily wages, weighted by days worked in each month. Male wage workers may not be representative of all males in the age group; indeed it is possible that *As* affects participation in the wage labor market. To obtain estimates IV estimates of the effects of *As* concentrations on wage rates that are not biased by sample selectivity, we employ a control function approach (Heckman, 1979), estimating a probit equation determining the probability that the

respondent earns a wage during the year. We assume that the decision to work maybe affected by the *As* lineage instrument (*As* concentrations being endogenous) as well as the value of owned landholdings and the number, age and gender of other adults in the household. The latter variables, however, accounting for their effect on selectivity, should not affect a worker's wage offer.

The first column of Table 7 reports the probit estimates of the probability of wage work for males aged 15-29. The estimates indicate that males with a lower genetic ability to excrete *As* (higher lineage *As* value) and thus with a higher concentrations of *As* do not participate more or less than other workers, but males from households with more valuable landholdings are significantly less likely to work. The second column of Table 7 reports the IV estimates for the sample of wage workers without correction for selectivity, using again the lineage instrument.[8] The last column includes the Mills ratio variable constructed from the probit estimates and thus corrects for selectivity. As can be seen, while the coefficient on the Mills ratio is marginally statistically significant, the estimate of the impact of arsenic on the wage is little affected in precision or magnitude. The point estimates indicate that a one standard deviation increase in the concentration of *As* reduces the male wage by 10%.

6. Are the Effects of *As* Concentrations Non-linear?

All of the specifications employed are either linear or log-linear, and the illustrative calculations of the effects of *As* concentrations on outcomes have used sample means, as is standard practice. Most arsenic abatement interventions that supplement diets to enhance the excretion of arsenic through methylation have marginal impacts on arsenic retention. But the impacts of arsenic concentrations may be non-linear. An important policy question therefore is which population groups should be targeted for interventions - those with the highest levels of *As* concentrations, or those with moderate levels (within the Bangladesh populations)? To help answer the question we obtained locally-weighted IV (Lowess) estimates of *As* concentration effects by levels of *As* contamination for all of the outcomes.

Figure 4 displays the IV estimates of the local effects of changes in *As* concentrations on the Raven's test score by level of arsenic concentration for men and women along with the associated confidence intervals of the estimates. The graph indicates clearly that the effects of *As* concentrations on women's performance are substantially smaller than those on men's and are also uniform across the distribution of *As* concentration levels. In contrast, the marginal effect of increasing *As* concentrations for men are larger in absolute value and negative at every level, but the effects are smaller at higher levels of *As* contamination. The distributions of *As* concentrations in the sample indicate that the median concentration is approximately 700 ppb. Elimination of all arsenic contamination down to levels such as those found in the US sample according to the non-linear IV estimates for the median male would increase the number of correct answers in the abridged Raven's CPM by

[8]The Sargan overidentification test statistic cannot reject the null hypothesis that the excluded first-stage variables are uncorrelated with the log wage residuals. χ^2 (5) =3.02, p=.696.

one-third, or by 6.3%. For men with concentrations at the 80th percentile, the same reduction in *As* would increase test performance, using the same mean, by 5%, and for the top 10%, by only 3.8%.

Figures 5 through 9 display the local IV estimates by *As* concentration levels for the other outcome measures where we found statistically significant results. In all cases, the effects of marginal changes in arsenic concentrations have larger effects on outcomes at lower levels of arsenic concentrations. Marginal improvements in lowering *As* ingestion or in enhancing excretion have bigger impacts on performance, earnings and nutritional status for groups with lower, but still high, levels of retained *As*. For example, the log wage estimates in Figure 10 indicate that eliminating all arsenic concentrations for the median male wage worker would increase the wage rate by 8.8%, while the same marginal 700 ppb reduction for workers in the top 10% of the *As* distribution would only increase the wage by 3.5%.

7. Will Economic Growth Alone Reduce Arsenic Concentration Levels?

Increases in income alone will not eliminate arsenic in the environment. However, one of the important mechanisms by which arsenic is ingested and methalyzed where water sources are contaminated by arsenic is through the consumption of vegetables and fruits. Our estimates in Table 3 suggest that these food groups increase retained arsenic levels on average. Income growth, by changing the composition of diets, could thus affect the levels of retained *As* in the population. In this section, we look at the relationship between the consumption of foods and land value to gain insights into how higher wealth or income levels affect diet and ultimately *As* concentrations.

Table 8 presents village fixed-effects estimates of the relationship between a household's value of land and the individual log consumption of green vegetables, all other vegetables, fruits, grains and tubers. Also included in the specification is the individual's age, age squared and gender. In this log-linear specification, for all but (non-green) vegetables, there is no statistically significant relationship between land wealth and consumption. The result for vegetables indicates that increases in income will increase vegetable consumption, which our estimates in Table 3 suggest would increase *As* concentrations. The point estimate indicates, however, that the income effect is small - a 100,000 *taka* (US$1400) increase in wealth is associated with only a 1.7% increase in the consumption of non-green vegetables.

Wealth effects may also be non-linear. Figures 11-13 display the locally-weighted (Lowess) FE estimates of the coefficients for the three food groups that we found to have non-zero effects on arsenic concentration levels - green vegetables, other vegetables, and fruits - by wealth level. The estimates indicate that wealth and the consumption of fruits and non-green vegetables are positively related to wealth at all wealth levels, with the positive wealth effects larger at lower wealth levels where the bulk of the Bangladesh population is situated. The differences in wealth effects by wealth level are not large. For example, at the wealth median (174,000 *taka*) a 100,000 *taka* increase in wealth would increase fruit and non-green vegetable consumption by 2.5%; at the 90th percentile (960,000 *taka*), the same increase in wealth increases consumption

of this food group by 2.3%. For fruit, the corresponding effects are 1.9% and 1.7%. For green vegetables, wealth effects are small, linear and negative. These estimates thus suggest that income growth alone will have little effect on *As* concentrations in the Bangladesh population, as changes in the composition of foods consumed is a major mechanism through which incomes can affect arsenic ingestion, and wealth effects are small and of mixed sign on the food groups that affect *As* concentrations. Consistent with this, the locally-weighted FE estimates of the relationship between land wealth and *As,* displayed in Figure 14*,* indicate that wealth effects are not statistically significantly different from zero at every wealth level.

8. Conclusion

In this paper we have used evidence on the genetic basis of arsenic excretion and unique information on family links among respondents living in different environments from a large panel survey to obtain new estimates of the effects of variation in arsenic concentrations retained in individuals on their cognitive and physical capabilities and earnings. Based on toenail clipping biomarkers, we found that the level of arsenic concentrations in the rural Bangladesh population are over 18 times those measured in subjects residing in the United States using the same laboratory technology, vary considerably across individuals but are no more elevated in high- than in low-wealth or high- than in low-educated sub-populations. We also found that, consistent with genetic models, *As* concentrations net of the influence of dietary intakes, were strongly correlated across immediate family members living apart. Based on methods that exploit these associations, we found that increases in *As* concentrations significantly reduce cognitive ability among men, but not women, reduce weight and weight-for-height and school attendance for both men and women, and reduce male wage rates. We also found that the effects of *As* concentrations are smaller the higher the concentration of arsenic, indicating that targeting moderately contaminated, rather than highly-contaminated populations would increase the efficacy of ameliorative interventions in arsenic-contaminated environments. Finally, while we found that increases in the consumption of vegetables and fruits increased the concentration of *As* in respondents, in the given environment in Bangladesh, the association between wealth and diet composition was not strong. Our estimates suggest that income growth alone will not reduce the cognitive, physical and earnings losses caused by the contamination of water in Bangladesh.

References

Adomako EE, Solaiman AR, Williams PN, Deacon C, Rahman GK, Meharg AA. (2009) Enhanced transfer of arsenic to grain for Bangladesh grown rice compared to US and EU. Environment International 35(3):476-9

Ahmad K, Hassan N. (1982) Nutrition survey of rural Bangladesh 1981-1982 Institute of Nutrition and Food Science, University of Dhaka

Asadullah MN, Chaudhury N. (2008) Social Interactions and Student Achievement in a Developing Country: An Instrumental Variables Approach. World Bank Policy Research Working Paper

British Geological Survey. (1999). Groundwater Studies for Arsenic Contamination in Bangladesh Phase 1 British Geological Survey [http://www.bgs.ac.uk/arsenic/home.html]

Chen Y, Factor-Litvak P, Howe GR, Graziano JH, Brandt-Rauf P, Parvez F, van Geen A, Ahsan H. (2007) Arsenic exposure from drinking water, dietary intakes of B vitamins and folate, and risk of high blood pressure in Bangladesh: a population-based, cross-sectional study. American Journal Epidemiology 165(5):541-52

Gamble MV, Liu X, Ahsan H, Pilsner R, Ilievski V, Slavkovich V, Parvez F, Levy D, Factor-Litvak P, Graziano JH. (2005) Folate, homocysteine, and arsenic metabolism in arsenic-exposed individuals in Bangladesh. Environmental Health Perspectives 113(12):1683-8

Gamble MV, Liu X, Slavkovich V, Pilsner JR, Ilievski V, Factor-Litvak P, Levy D, Alam S, Islam M, Parvez F, Ahsan H, Graziano JH. (2007) Folic acid supplementation lowers blood arsenic. American Journal of Clinical Nutrition 86(4):1202-9

Guha Mazumder DN, Haque R, Ghosh N, De BK, Santra A, Chakraborty D, Smith AH. (1998) Arsenic levels in drinking water and the prevalence of skin lesions in West Bengal, India. International Journal Epidemiology 27(5):871-7

Heck JE, Gamble MV, Chen Y, Graziano JH, Slavkovich V, Parvez F, Baron JA, Howe GR, Ahsan H. (2007) Consumption of folate-related nutrients and metabolism of arsenic in Bangladesh. American Journal of Clinical Nutrition 85(5):1367-74

Heckman JJ (1979) Sample Selection Bias as a Specification Error. Econometrica 47 (1): 153-161.

Hossain M.F. (2006) Arsenic contamination in Bangladesh-An overview. Agriculture, Ecosystems and Environment 113(1-4): 1-16

Kelley, K.A., Plank, T., Ludden, J.N., Staudigel, H., 2003. Composition of altered oceanic crust at ODP Sites 801 and 1149. Geochem. Geophys. Geosys. 4, doi:10.1029/2002GC000435.

Kile ML, Houseman EA, Breton CV, Smith T, Quamruzzaman Q, Rahman M, Mahiuddin G, Christiani DC. (2007). Dietary arsenic exposure in Bangladesh. Environmental Health Perspectives 115(6):889-93

Lindberg AL, Kumar R, Goessler W, Thirumaran R, Gurzau E, Koppova K, Rudnai P, Leonardi G, Fletcher T, Vahter M. (2007). Metabolism of low-dose inorganic arsenic in a central European population: influence of sex and genetic polymorphisms. Environmental Health Perspectives 115(7):1081-6

Lindberg AL, Ekström EC, Nermell B, Rahman M, Lönnerdal B, Persson LA, Vahter M. (2008) Gender and age differences in the metabolism of inorganic arsenic in a highly exposed population in Bangladesh. 106(1):110-20

Mitra SR, Mazumder DN, Basu A, Block G, Haque R, Samanta S, Ghosh N, Smith MM, von Ehrenstein OS, Smith AH. (2004). Nutritional factors and susceptibility to arsenic-caused skin lesions in West Bengal, India. Environmental Health Perspectives 112(10):1104-9

Pitt MP, Rosenzweig MR, Hassan, NM (1990). Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries. The American Economic Review 80(5): 1139-1156.

Pitt MP, Rosenzweig MR, Hassan, NM (forthcoming). Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy.

Rosado JL, Ronquillo D, Kordas K, Rojas O, Alatorre J, Lopez P, Garcia-Vargas G, Del Carmen Caamaño M, Cebrián ME, Stoltzfus RJ. (2007). Arsenic exposure and cognitive performance in Mexican schoolchildren. Environmental Health Perspectives 115(9):1371-5

Roychowdhury T, Tokunaga H, Ando M. (2003). Survey of arsenic and other heavy metals in food composites and drinking water and estimation of dietary intake by the villagers from an arsenic-affected area of West Bengal, India. Science of the Total Environment 308(1-3):15-35

Smith AH, Lingas EO, Rahman M. (2000). Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. Bulletin of the World Health Organization 78(9):1093-103

Steinmaus C, Carrigan K, Kalman D, Atallah R, Yuan Y, Smith AH. (2005). Dietary intake and arsenic methylation in a U.S. population. Environmental Health Perspectives 113(9):1153-9

Styblo M, Thomas DJ. (2001). Selenium modifies the metabolism and toxicity of arsenic in primary rat hepatocytes. Toxicology and Applied Pharmacology 172(1):52-61

Subcommittee on Arsenic in Drinking Water. (1999). Arsenic in Drinking Water Report. National Research Council

Subcommittee to Update the 1999 Arsenic in Drinking Water Report. (2001). Arsenic in Drinking Water: 2001 Update. National Research Council

Tice RR, Yager JW, Andrews P, Crecelius E. (1997). Effect of hepatic methyl donor status on urinary excretion and DNA damage in B6C3F1 mice treated with sodium arsenite. Mutation Research 386(3):315-34

Uchino T, Roychowdhury T, Ando M, Tokunaga H. (2006) Intake of arsenic from water, food composites and excretion through urine, hair from a studied population in West Bengal, India. Food and Chemical Toxicology 44(4):455-61

US Environmental Protection Agency. (2001). National Primary Drinking Water Regulations; Arsenic and Clarifications to Compliance and New Source Contaminants Monitoring. Federal Register Environmental Documents 66(14): 6975-7066

Vahter M, Marafante E. (1987) Effects of low dietary intake of methionine, choline or proteins on the biotransformation of arsenite in the rabbit. Toxicology Letters 37(1):41-6

Vahter M. (2000) Genetic polymorphism in the biotransformation of inorganic arsenic and its role in toxicity. Toxicology Letters 112-113:209-17

von Ehrenstein OS, Mazumder DN, Yuan Y, Samanta S, Balmes J, Sil A, Ghosh N, Hira-Smith M, Haque R, Purushothamam R, Lahiri S, Das S, Smith AH. (2005) Decrements in lung function related to arsenic in drinking water in West Bengal, India. American Journal of Epidemiology 162(6):533-41

Wasserman GA, Liu X, Parvez F, Ahsan H, Factor-Litvak P, van Geen A, Slavkovich V, LoIacono NJ, Cheng Z, Hussain I, Momotaj H, Graziano JH. (2004). Water arsenic exposure and children's intellectual function in Araihazar, Bangladesh. Environmental Health Perspectives 112(13):1329-33

Wasserman GA, Liu X, Parvez F, Ahsan H, Factor-Litvak P, Kline J, van Geen A, Slavkovich V, LoIacono NJ, Levy D, Cheng Z, Graziano JH. (2007) Water Arsenic Exposure and Intellectual Function in 6-Year-Old Children in Araihazar, Bangladesh. Environmental Health Perspectives 115(2): 285-289

Watanabe C, Matsui T, Inaoka T, Kadono T, Miyazaki K, Bae MJ, Ono T, Ohtsuka R, Bokul AT. (2007) Dermatological and nutritional/growth effects among children living in arsenic-contaminated communities in rural Bangladesh. Journal of Environmental Science and Health, Part A 42(12):1835-41

40. Zablotska LB, Chen Y, Graziano JH, Parvez F, van Geen A, Howe GR, Ahsan H. (2008) Protective effects of B vitamins and antioxidants on the risk of arsenic-related skin lesions in Bangladesh. Environmental Health Perspectives 116(8):1056-62

41. Zavala YJ, Duxbury JM. (2008). Arsenic in rice: I. Estimating normal levels of total arsenic in rice grain. Environmental Science and Technology 42(10):3856-60

42. Zhu Y-G, Williams PN, Meharg AA. (2008), Exposure to inorganic arsenic from rice: A global health issue? Environmental Pollution 154: 169-171

Appendix A

Analytical Methods For Determination of *As* in Toenail Solutions

Solutions were analyzed for *As* concentration using a Thermo X-Series II quadrupole inductively-coupled plasma mass spectrometer (ICP-MS) equipped with collision cell technology (CCT) at the Graduate School of Oceanography, University of Rhode Island. The CCT was flushed with a *He-H* gas mix to break up the $^{40}Ar^{35}Cl$ interference on ^{75}As. Daily assessment of the efficiency of the collision cell, monitoring mass 75 in an ultra-pure 2% HCl solution , showed that the CCT reduced the interference to <0.02% $^{40}Ar^{35}Cl/^{35}Cl$. In practical terms, this translated to count rates of <1 count per second attributed to *Cl* in the unknown solutions, effectively removing the interference. All samples were run using identical analysis and data reduction protocols, similar to procedures described by Kelley *et al.* (2003). A procedural blank was subtracted from all raw data, and count rates were corrected for instrumental drift using an external drift correcting solution analyzed every 5 samples, which was matrix-matched to the unknown samples and spiked with 15 ppb *As*. Concentrations of *As* in unknown solutions were determined by calibrating against four single-element *As* solution standards spanning a range from 0.5 to 100 ppb. Calibration curves were linear, with r>0.9990 for all analytical runs. Arsenic concentrations in unknown solutions were then corrected for the dilution of each starting sample weight, to yield the concentration of *As* in the raw toenail material. Each unknown solution was analyzed twice in every analytical session, and within-session reproducibility was <10% rsd for these replicate analyses (on average, within 5% rsd). Some samples contained enough material to perform replicate dissolutions of the starting material (n=184), and these replicate dissolutions are in excellent agreement, on average within 15% rsd. Lower precision for replicate dissolutions of the raw material is not surprising, likely resulting from heterogeneities within the toenails, which were dissolved as whole clippings without attempting to homogenize the material in advance.

Figure 1. Distribution of *As* Concentrations (ppb), by Sample

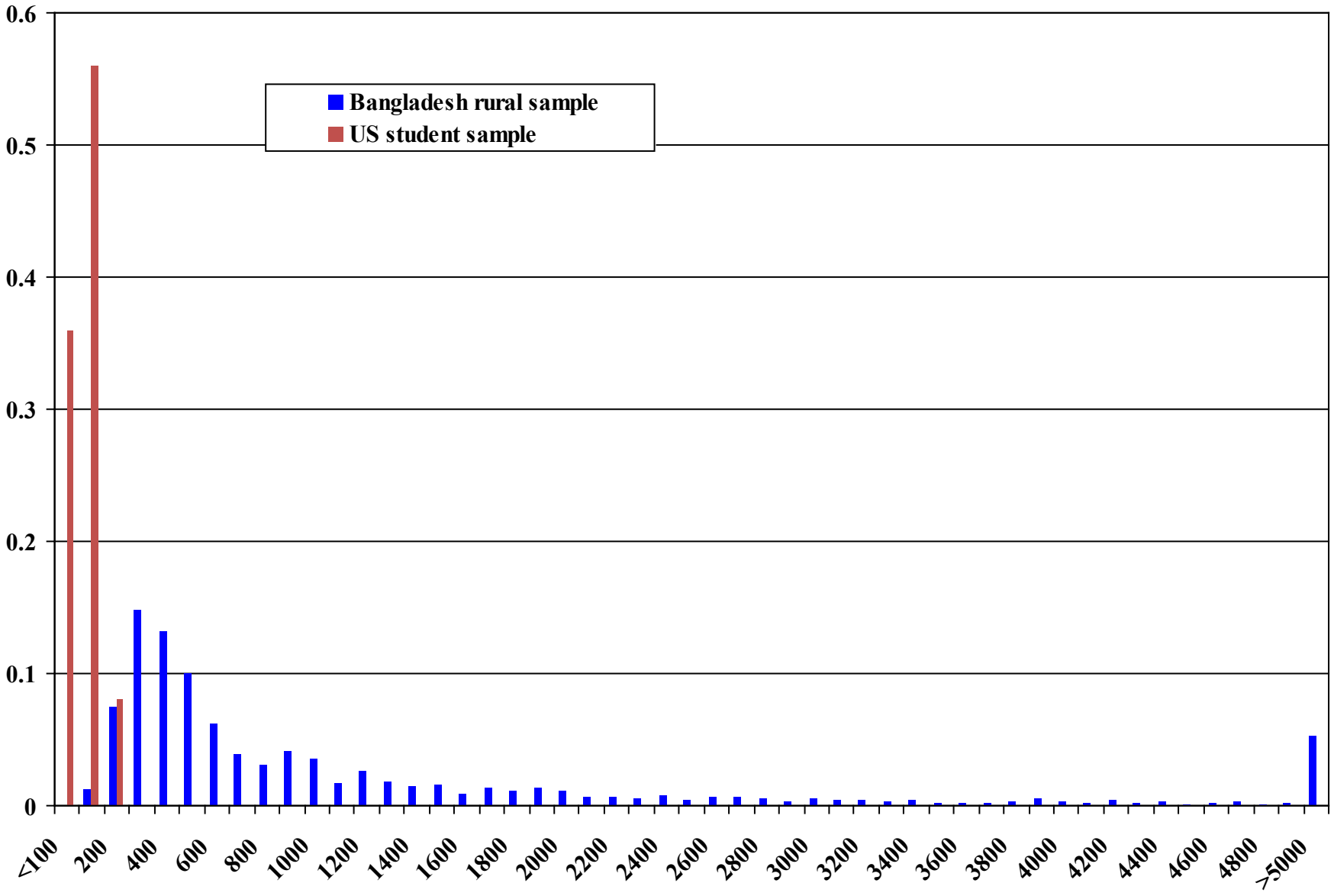

Figure 2. Mean *As* Concentrations (ppb) by Owned Landholdings (square decimeters)

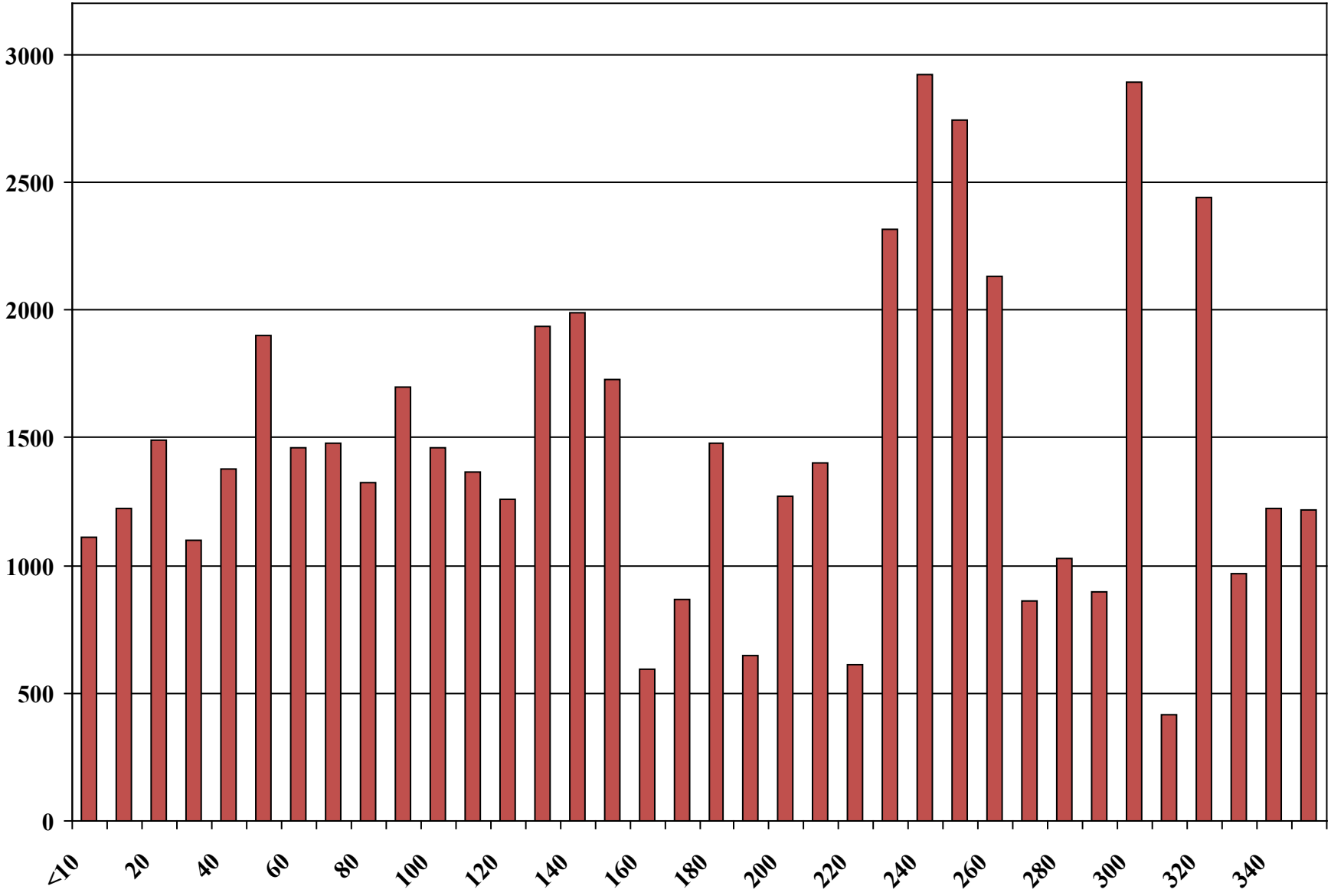

Figure 3. Mean *As* Concentration (ppb) for Respondents Aged 20-59 by Years of Schooling

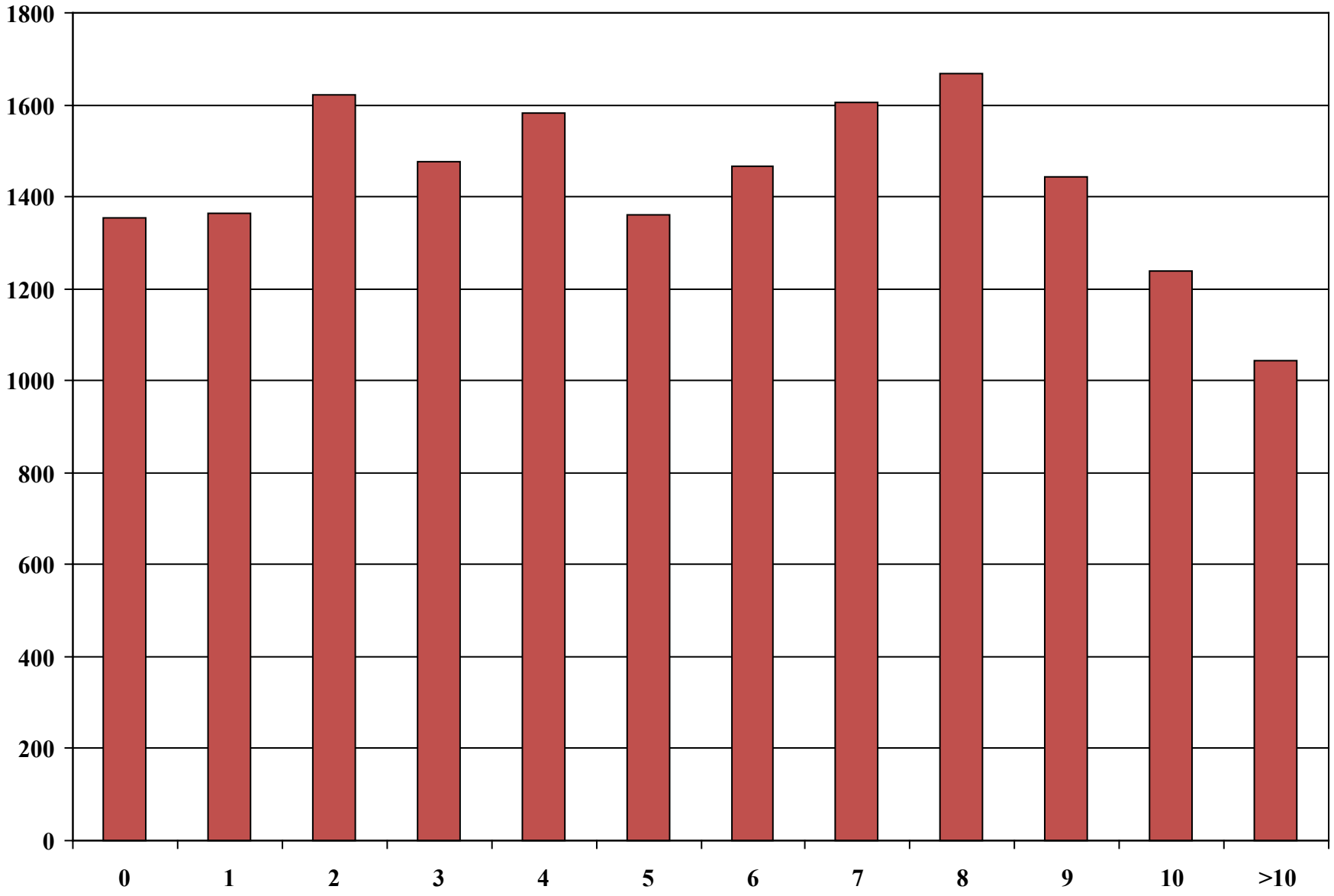

Figure 4. Locally-weighted IV Estimates of the Effects of *As* Concentrations (ppb) on Abridged Raven's CPM Test Performance, by Gender and *As* Concentration

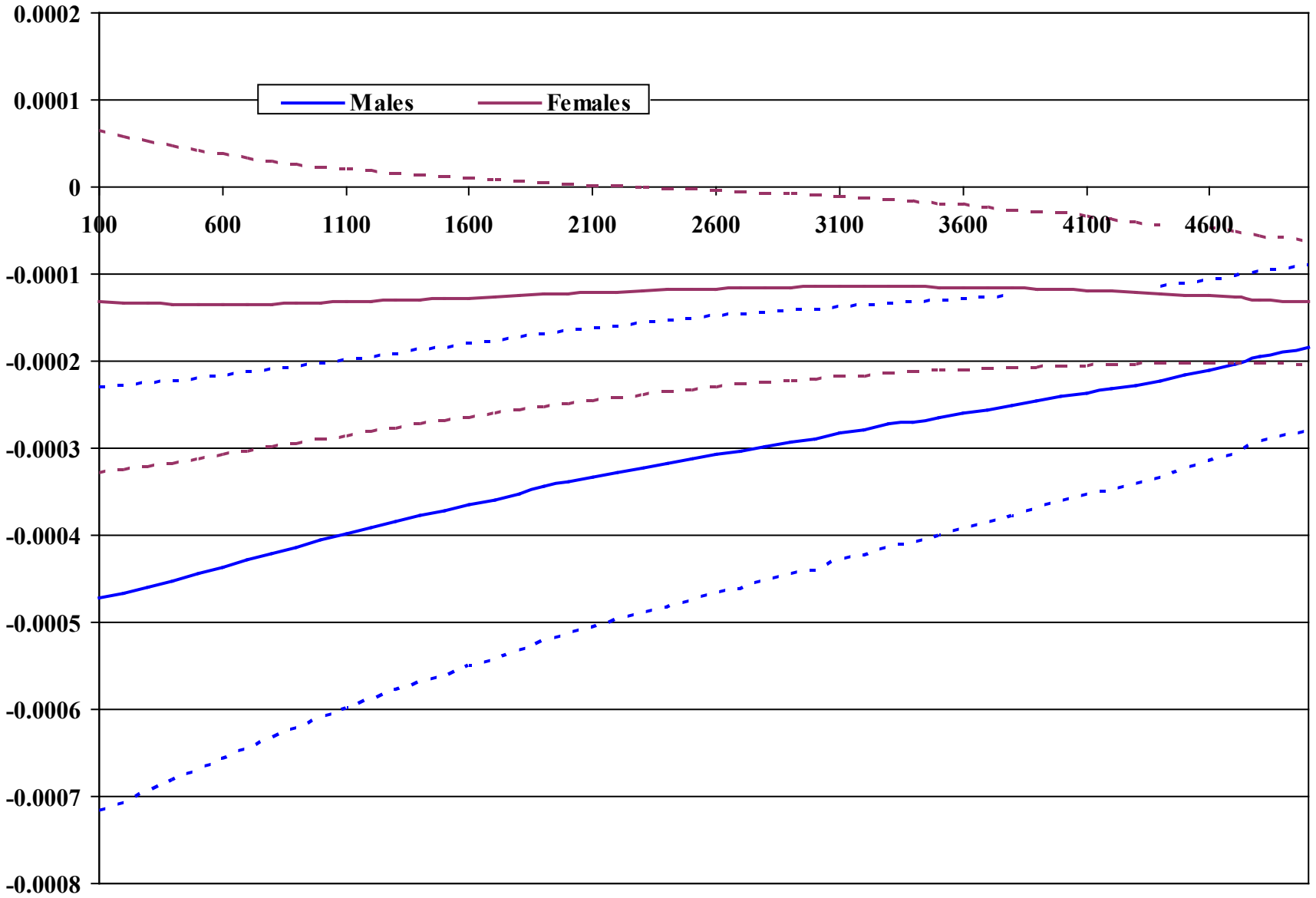

Figure 5. Locally-weighted IV Estimates of the Effects of *As* Concentrations (ppb) on Male Pinch Strength, by *As* Concentration

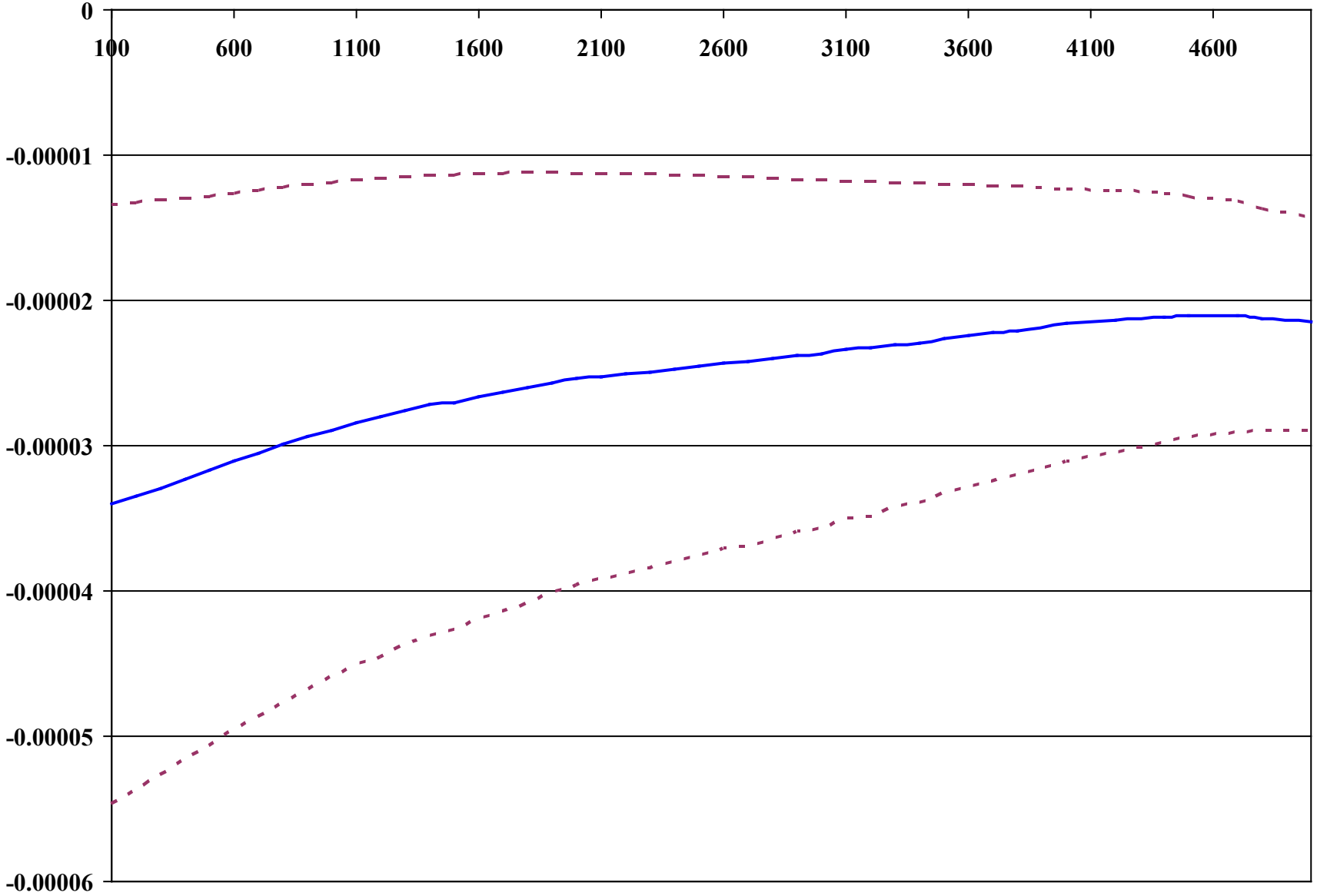

Figure 6. Locally-weighted IV Estimates of the Effects of *As* Concentrations (ppb) on Male Grip Strength, by *As* Concentration

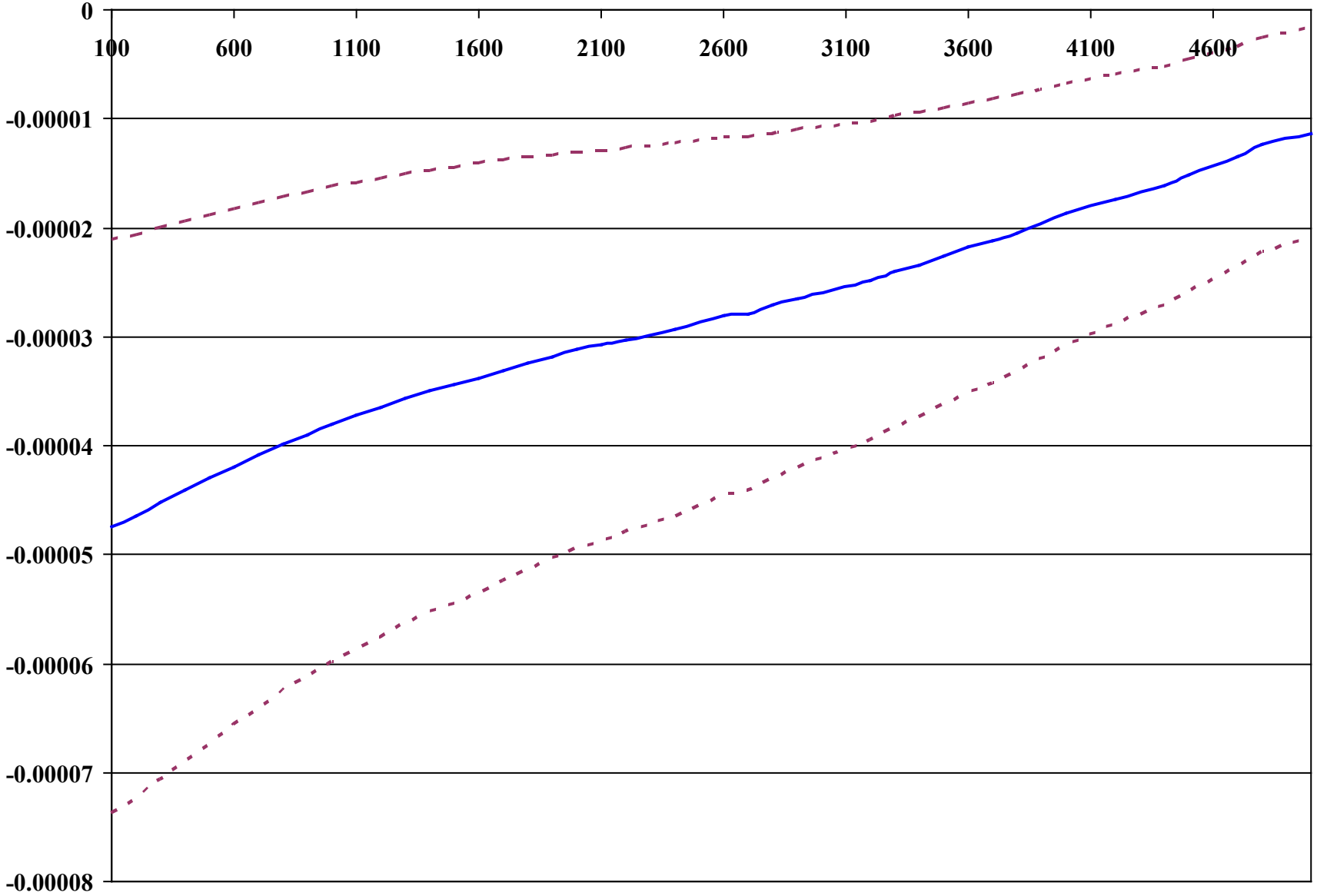

Figure 7. Locally-weighted IV Estimates of the Effects of *As* Concentrations (ppb) on Weight, by *As* Concentration

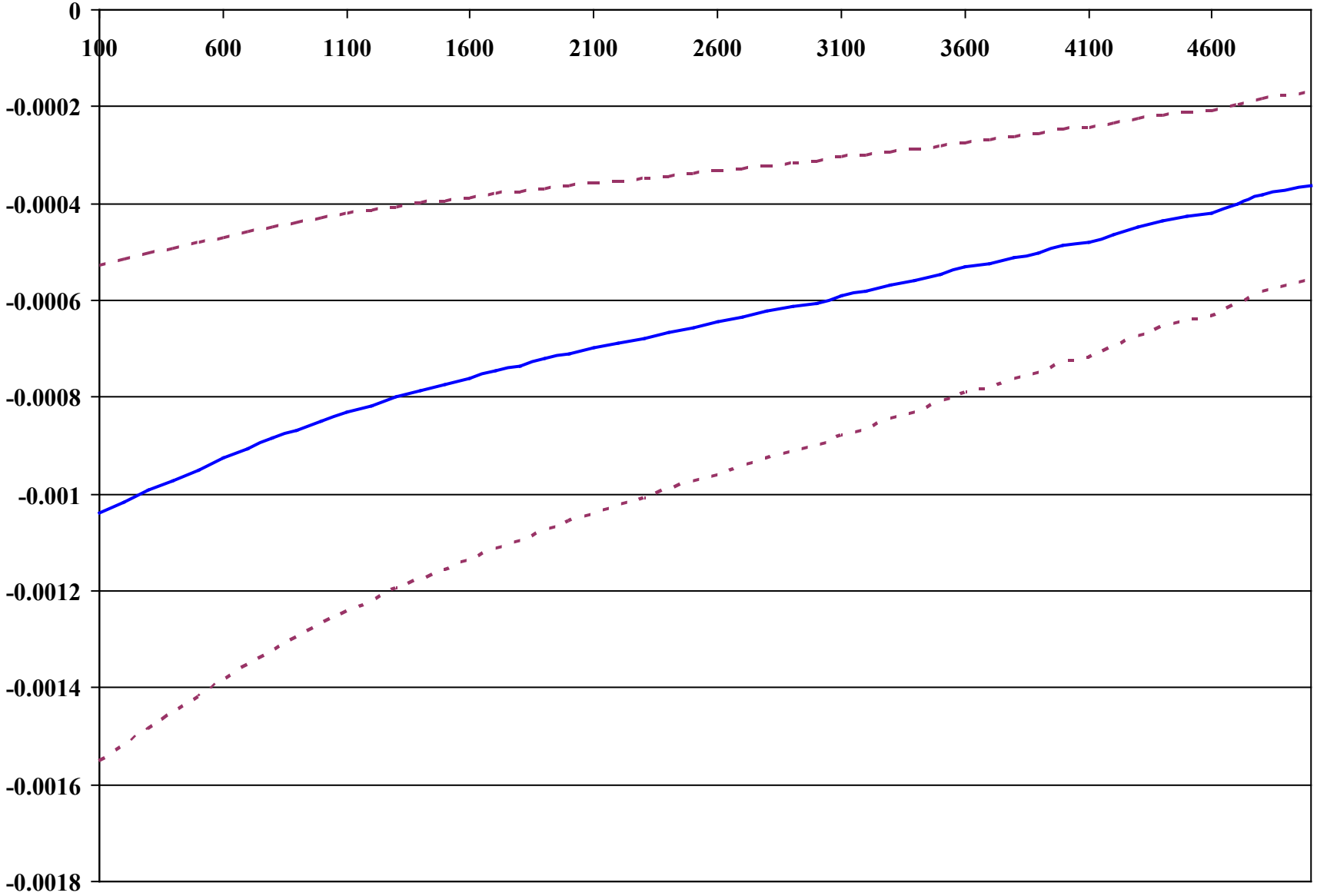

Figure 8. Locally-weighted IV Estimates of the Effects of *As* Concentrations (ppb) on Weight for Height, by *As* Concentration

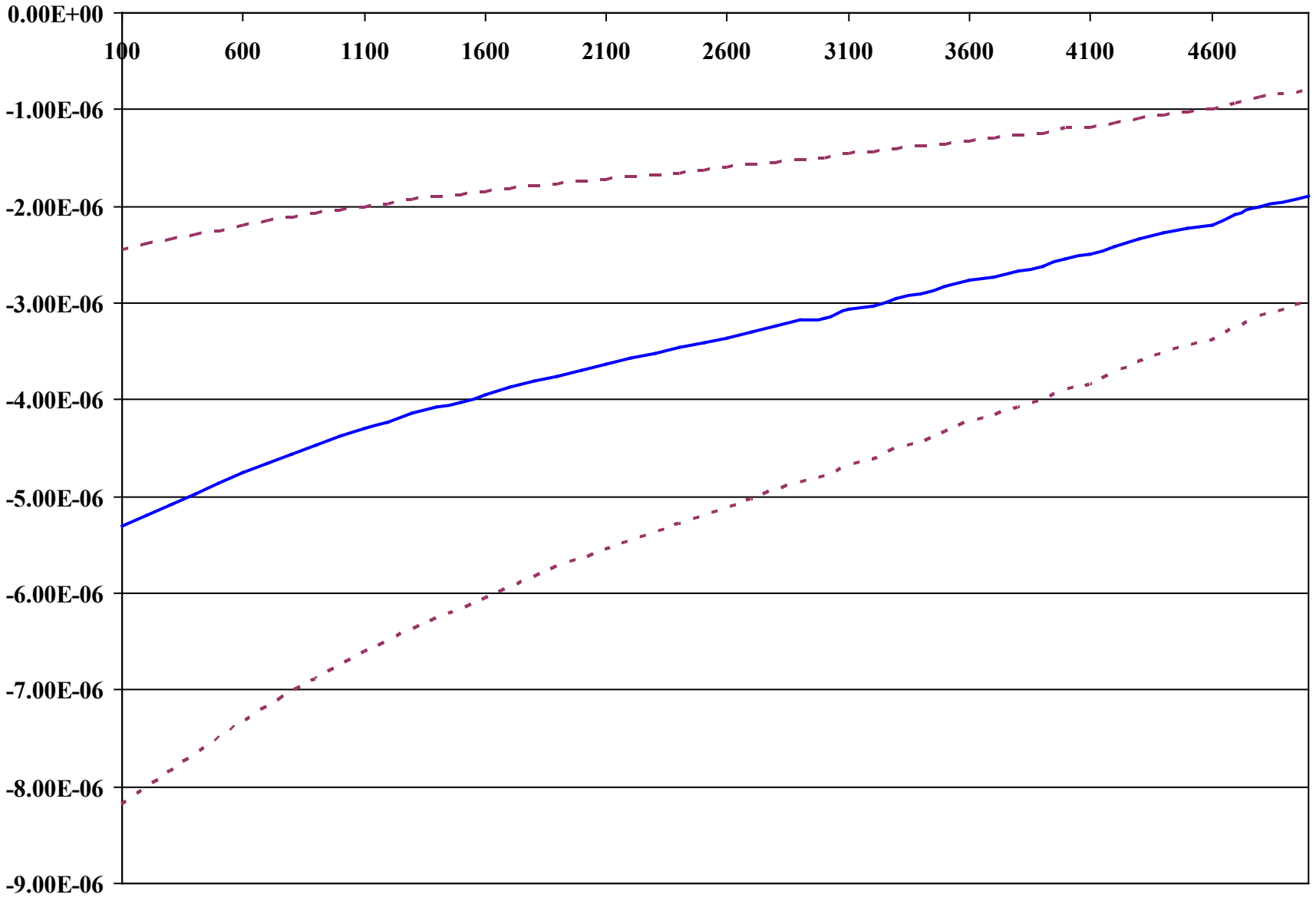

Figure 9. Locally-weighted IV Estimates of the Effects of *As* Concentrations (ppb) on the Probability of Attending School, by *As* Concentration

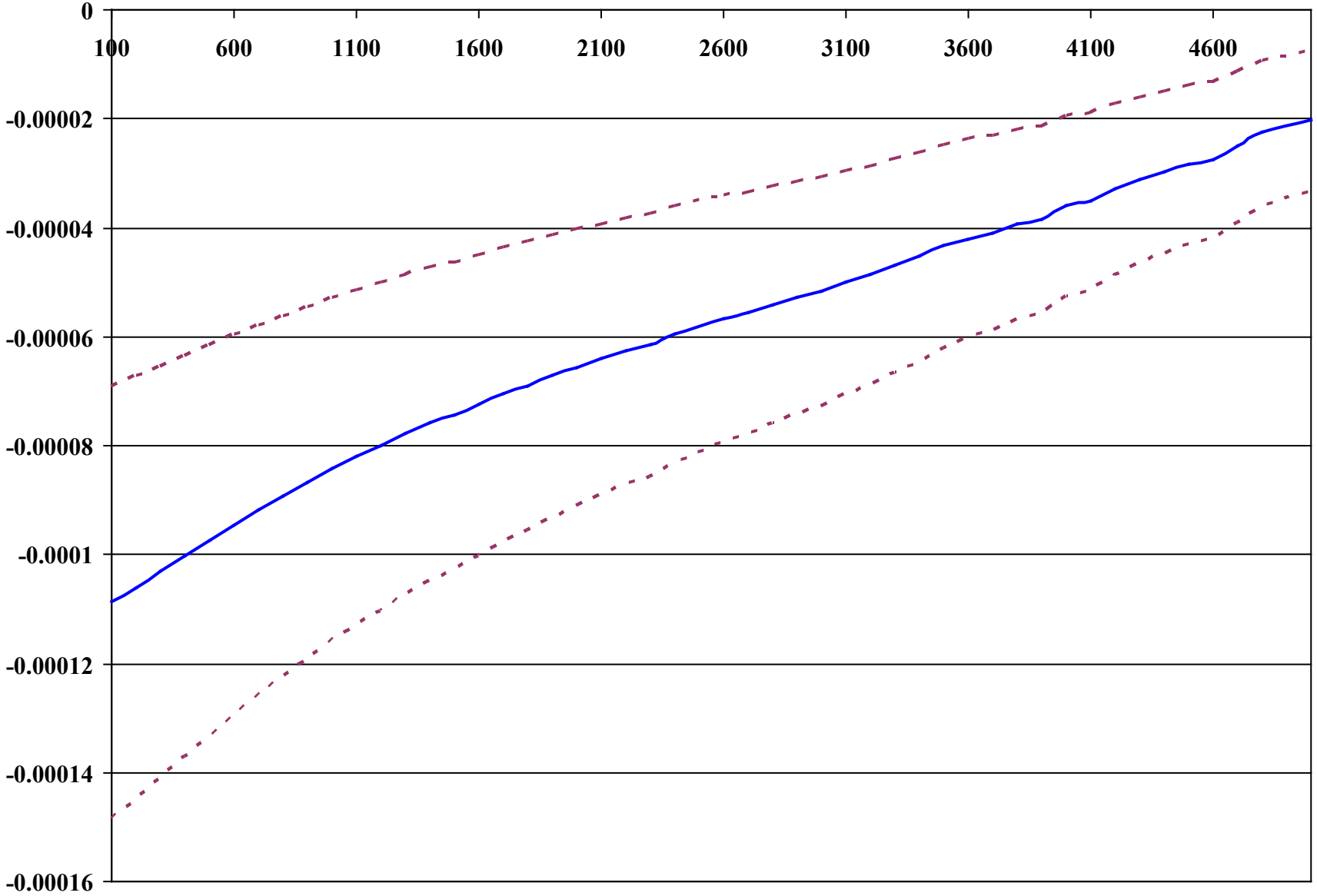

Figure 10. Locally-weighted IV Estimates of the Effects of *As* Concentrations (ppb) on the Log Daily Male Wage, by *As* Concentration

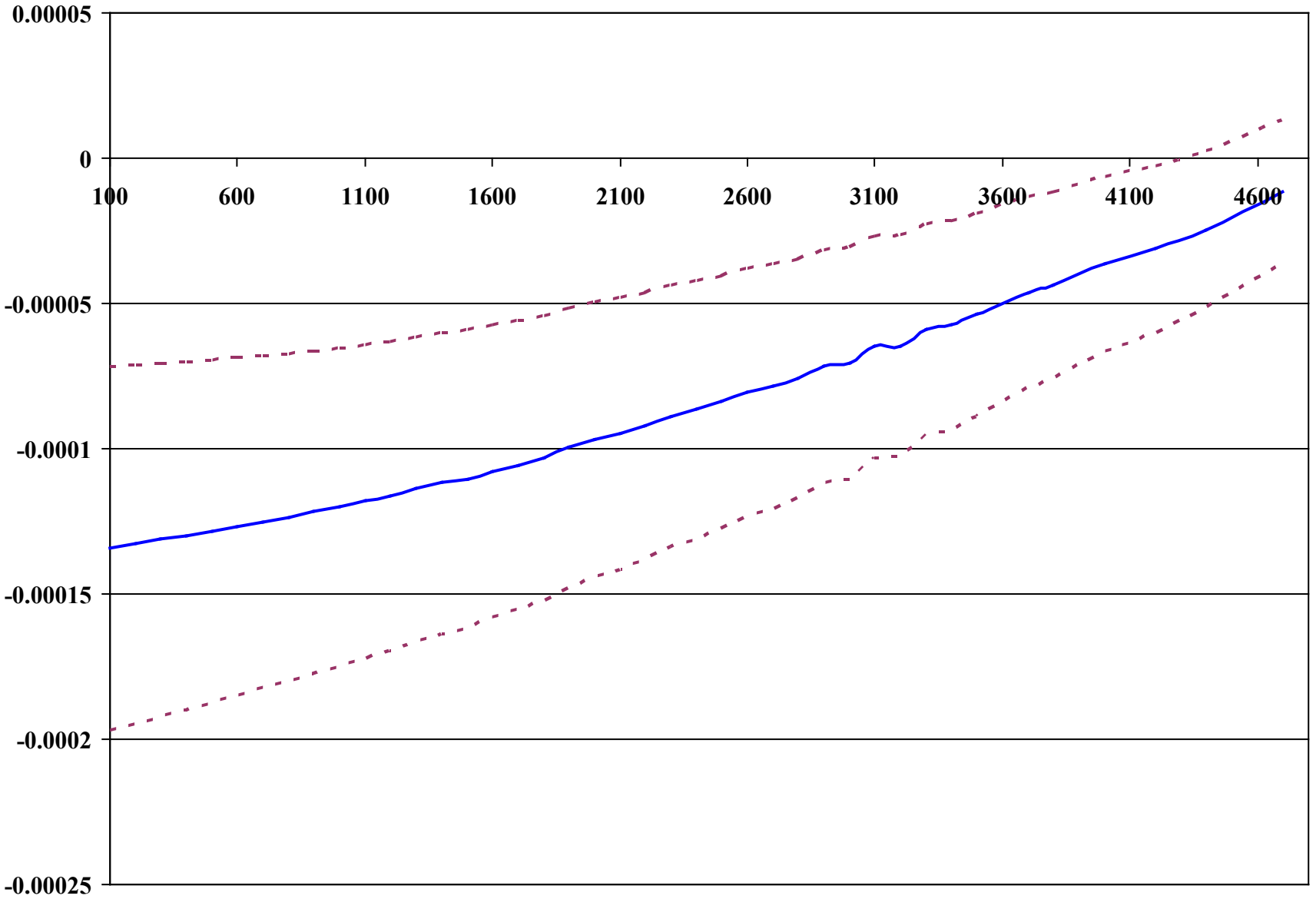

Figure 11. Locally-weighted FE Estimates of the Effects Land Value on Log Vegetables, by Land Value

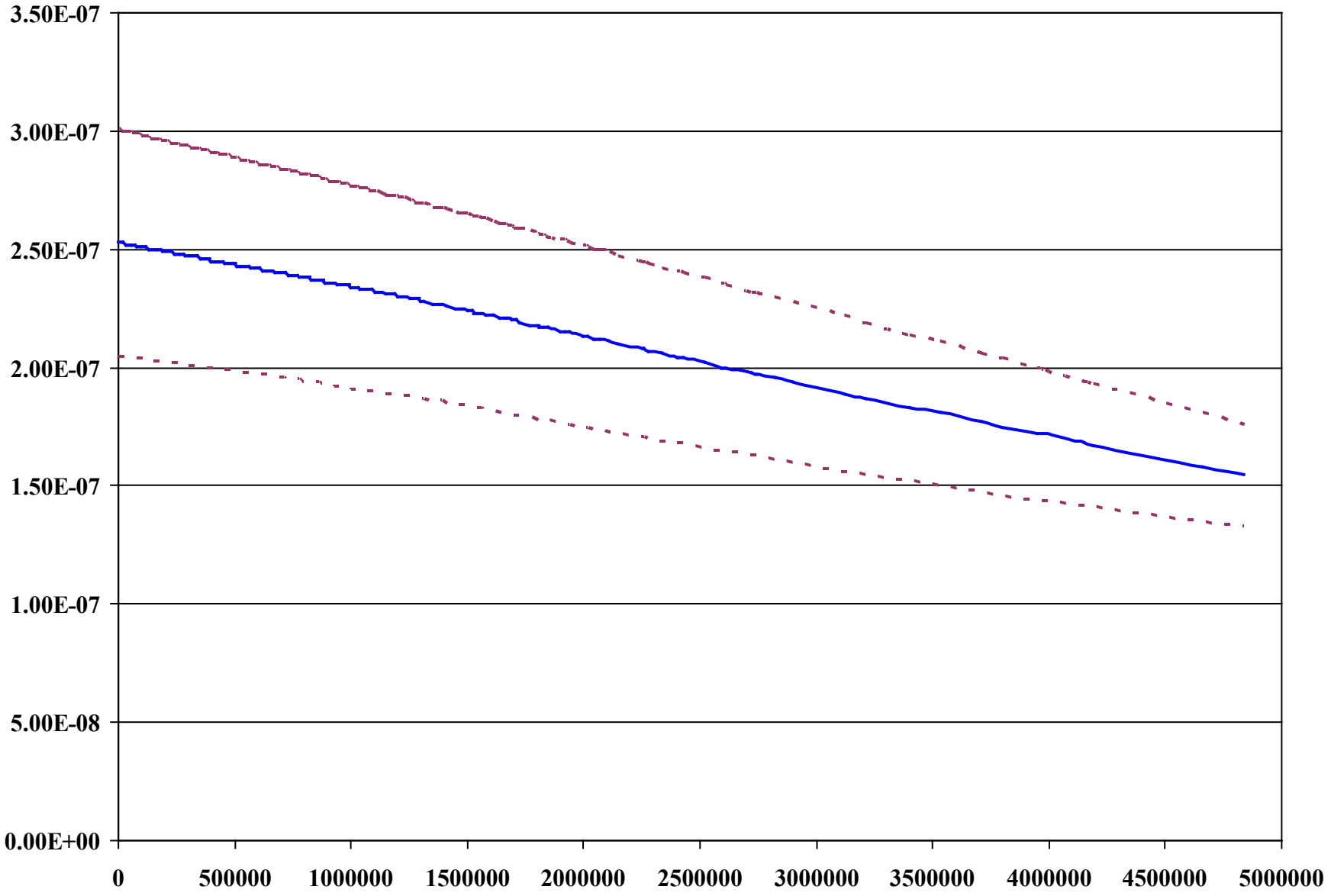

Figure 12. Locally-weighted FE Estimates of the Effects Land Value on Log Fruits, by Land Value

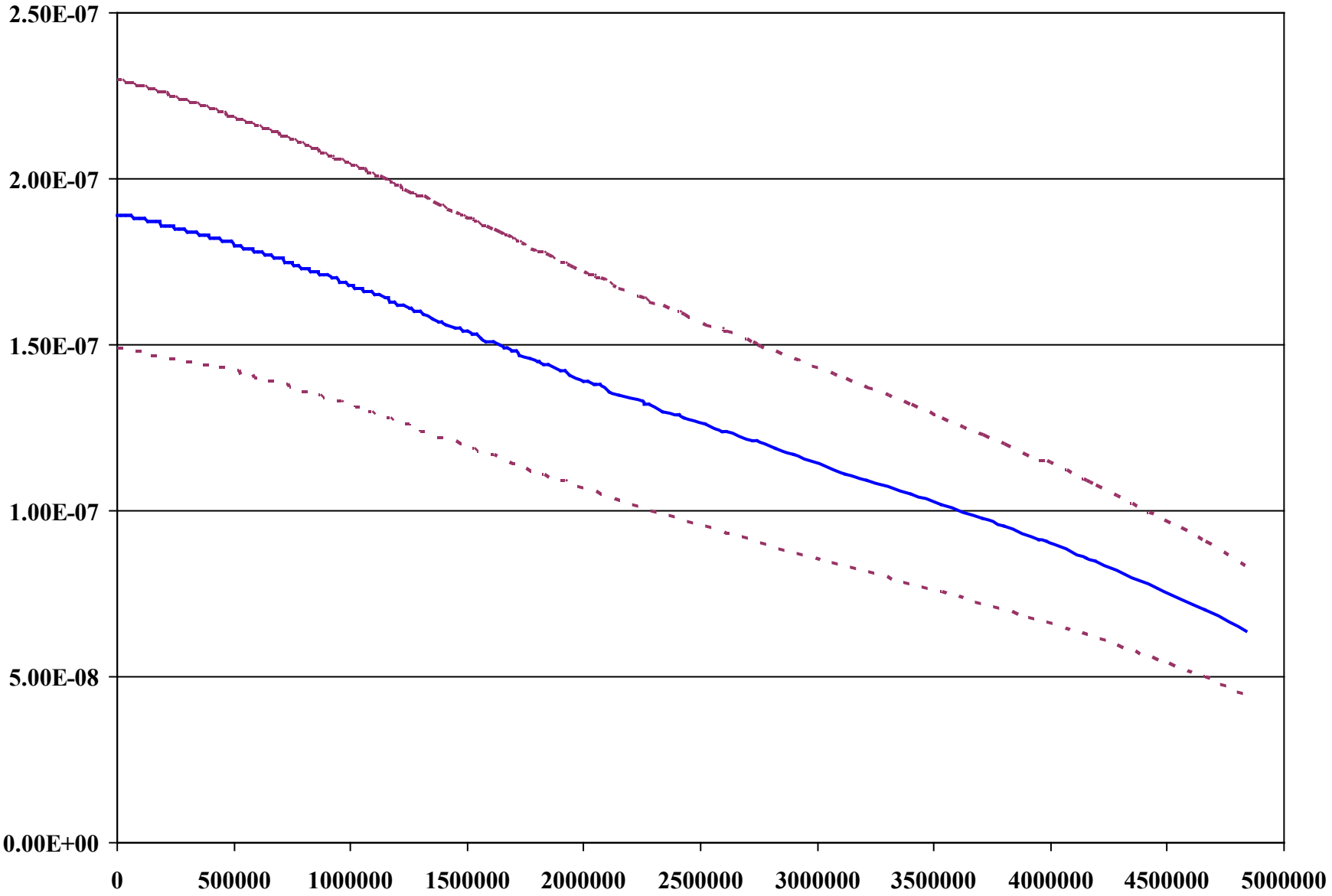

Figure 13. Locally-weighted FE Estimates of the Effects Land Value on Log Green Vegetables, by Land Value

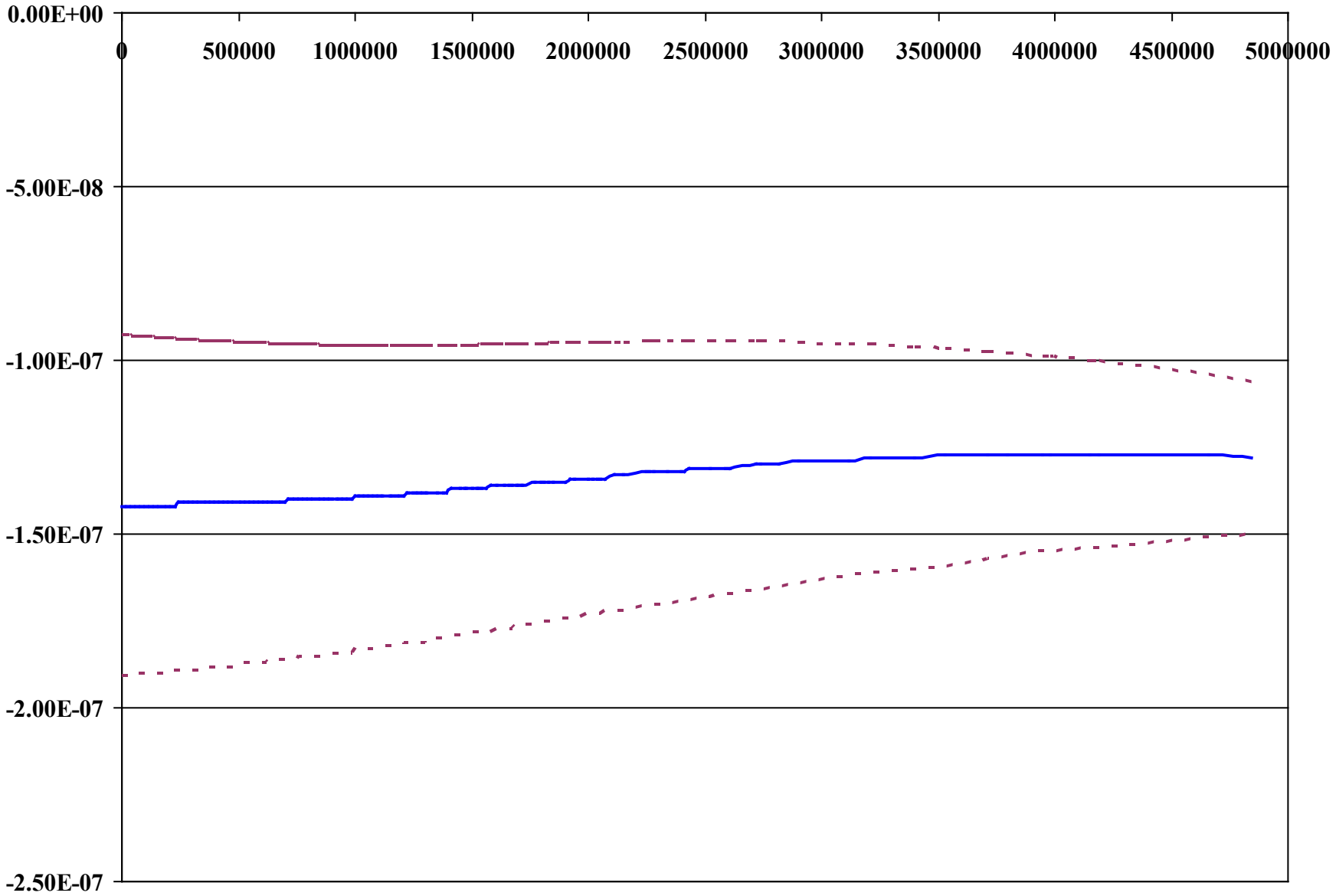

Figure 14. Locally-weighted FE Estimates of the Effects Land Value on *As* Concentrations, by Land Value

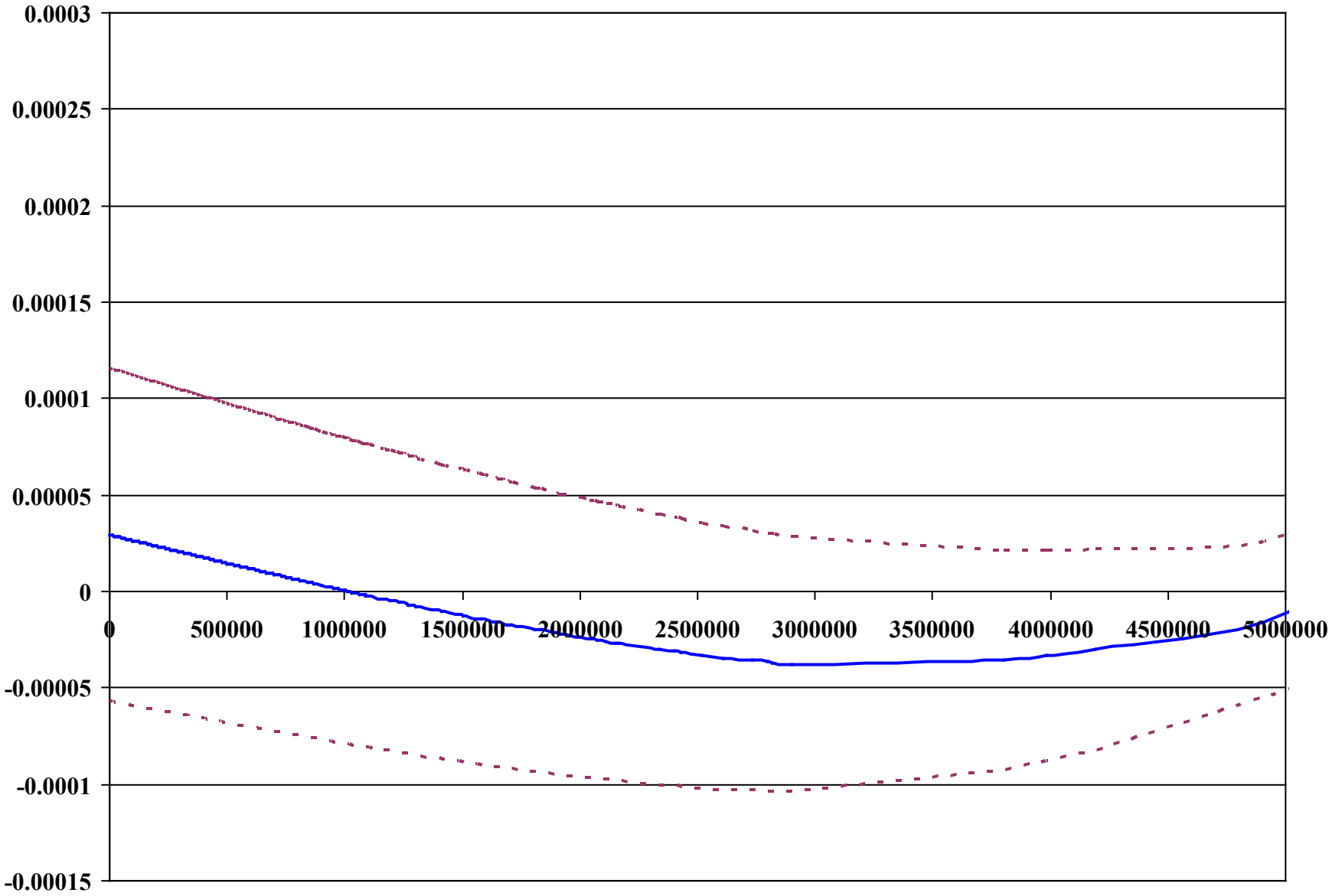

Table 1
As Concentration and Diet, by Gender: Respondents Aged 15-29

	All	Males	Females
As (ppb x 10^{-3})	1294.1*	1205.7	1383.6
	(1860.1)	(1733.5)	(1977.6)
Grain consumption	433.3	435.1	431.6
	(218.1)	(270.0)	(148.1)
Pulse consumption	12.9	13.4	12.4
	(26.8)	(28.4)	(25.1)
Green vegetables consumption	31.5*	27.8	35.3
	(72.7)	(70.6)	(74.6)
Other vegetable consumption	102.2	104.3	100.0
	(132.5)	(143.9)	(119.8)
Tuber consumption	84.1*	79.1	89.1
	(89.8)	(92.8)	(86.5)
Fruit consumption	13.9	14.2	13.7
	(55.6)	(64.4)	(45.1)
Meat consumption	67.7	68.8	66.6
	(88.9)	(96.3)	(80.8)
N	1170	589	581

Standard deviations in parentheses. *Gender difference statistically significant at the .05 level, one-tailed test.

Table 2
Outcome Measures, by Gender: Respondents Aged 15-29

	All	Males	Females
Abridged Raven's CPM (correct answers)	4.53*	4.98	4.08
	(2.09)	(2.11)	(1.98)
Log pinch strength	3.55*	3.69	3.42
	(.240)	(.191)	(.164)
Log grip strength	4.76*	4.95	4.54
	(.304)	(.227)	(.220)
Height	156.2*	161.9	150.5
	(8.62)	(7.17)	(5.70)
Weight	46.1*	48.7	43.4
	(7.92)	(7.91)	(7.00)
Weight/Height	.294*	.300	.288
	(.042)	(.042)	(.042)
Attending School (ages 15-23)	.0797*	.118	.0502
	(.271)	(.323)	(.218)
Years of school completed	6.11	5.96	6.27
	(3.48)	(3.64)	(3.30)
Daily wage	125.5*	94.8	128.0
	(84.8)	(53.1)	(86.5)

Standard deviations in parentheses. *Gender difference statistically significant at the .05 level, one-tailed test.

Table 3
As Production Function Estimates

Estimation procedure	OLS	IV
Grain consumption	-.00106	-.0911
	(0.01)	(0.52)
Pulse consumption	.0173	.0541
	(1.11)	(1.67)
Green vegetables consumption	.0166	.0667
	(1.15)	(2.21)
Other vegetable consumption	.0256	.0856
	(1.68)	(4.02)
Tuber consumption	-.0402	-.0906
	(1.35)	(1.56)
Fruit consumption	.0352	.0830
	(1.69)	(2.61)
Meat consumption	-.0152	-.00331
	(1.00)	(0.12)
Female	-.100	-.101
	(2.96)	(2.51)
Age	.000165	-.000096
	(0.15)	(0.08)
N	3971	3971
Anderson LM χ^2 (1) [*p*]	-	768.9
		[.000]

All variables in logs. Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients. The identifying instruments are village prices.

Table 4
IV Estimates of the Effects of *As* Concentrations on Cognition and Strength, by Gender:
Men and Women Aged 15-29

Measure	Abridged Raven's CPM			Log Pinch Strength			Log Grip Strength		
Group	All	Males	Females	All	Males	Females	All	Males	Females
As (ppb x 10^{-3})	-.155	-.259	-.0754	-.00845	-.0183	.00169	-.0121	-.0192	-.00369
	(2.64)	(2.55)	(1.08)	(1.73)	(2.20)	(0.29)	(1.63)	(1.69)	(0.39)
Value of household landholdings (x 10^{-6})	.380	.339	.491	.0126	.0117	.0133	.000444	.00501	-.00333
	(4.70)	(3.40)	(3.45)	(1.85)	(1.39)	(1.13)	(0.05)	(0.48)	(0.19)
Age	.210	.284	.0314	.0788	.138	.0102	.127	.188	.0372
	(1.25)	(1.18)	(0.13)	(5.60)	(6.87)	(0.52)	(6.42)	(7.30)	(1.21)
Age squared	-.00567	-0.00725	-.00191	-.00157	-.00280	-.00013	-.00257	-.00386	-.00061
	(1.48)	(1.33)	(0.35)	(4.86)	(6.13)	(0.28)	(5.69)	(6.52)	(0.86)
Female	-.963	-	-	-.273	-	-	-.408	-	-
	(7.22)			(24.2)			(23.2)		
N	1159	580	579	1168	588	580	867	453	414
Anderson LM χ^2 (1) [*p*]	344.2	139.3	203.9	352.8	147.3	204.4	259.2	150.9	144.8
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]

Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients. The specification also includes the number of adult men and women in the household over age 14 and their average age. The identifying instrument is the average *As* concentration residual for genetically-linked individuals not residing in the same village as the subject.

Table 5
IV Estimates of the Effects of *As* Concentrations on Nutritional Status, by Gender:
Men and Women Aged 15-29

Measure	Height (kg)			Weight (cm)			Weight/Height		
Group	All	Males	Females	All	Males	Females	All	Males	Females
As (ppb x 10^{-3})	-.344	-.556	-.129	-.512	-.395	-.529	-.00263	-.00143	-.00321
	(1.91)	(1.74)	(0.64)	(2.61)	(1.20)	(2.22)	(2.39)	(0.81)	(2.25)
Value of household landholdings (x 10^{-6})	.469	.329	.790	1.36	1.22	1.65	.00783	.00699	.00945
	(1.87)	(1.01)	(1.92)	(4.98)	(3.63)	(3.42)	(5.07)	(3.91)	(3.27)
Age	3.47	5.19	1.57	3.56	5.16	1.70	.0170	.0241	.00856
	(6.72)	(6.73)	(2.27)	(6.33)	(6.46)	(2.10)	(5.37)	(5.67)	(1.76)
Age squared	-.0728	-.109	-.0310	-.0676	-.102	-.0277	-.000312	-.000466	-.00013
	(6.16)	(6.22)	(1.95)	(5.25)	(5.61)	(1.48)	(4.29)	(4.81)	(1.16)
Female	-11.5	-	-	-5.41	-	-	-.0123	-	-
	(27.8)			(11.9)			(4.83)		
N	1170	589	581	1170	589	581	1170	589	581
Anderson LM χ^2 (1) [*p*]	353.5	147.6	204.7	353.5	147.6	204.7	353.5	147.6	204.7
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]

Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients. The specification also includes the number of adult men and women in the household over age 14 and their average age. The identifying instrument is the average *As* concentration residual for genetically-linked individuals not residing in the same village as the subject.

Table 6
IV Estimates of the Effects of *As* Concentrations on School Attendance for Men and Women Aged 15-23 and Schooling Attainment for Men and Women Aged 15-29

Measure	Attending School			Years of Schooling Completed		
Group	All	Males	Females	All	Males	Females
As (ppb x 10^{-3})	-.0337	-.0450	-.0274	-.0413	-.199	.0711
	(2.24)	(1.76)	(1.56)	(0.44)	(1.23)	(0.63)
Value of household landholdings ($x\ 10^{-6}$)	.0763	.0715	.0901	.943	.864	1.13
	(4.06)	(3.04)	(2.49)	(7.15)	(5.27)	(4.93)
Age	-.513	-.449	-.526	1.07	.804	1.31
	(5.21)	(2.88)	(4.24)	(3.96)	(2.06)	(3.41)
Age squared	.0117	.00994	.0122	-.0245	-.0173	-.0307
	(4.51)	(2.42)	(3.75)	(3.95)	(1.95)	(3.47)
Female	-.0947	-	-	.606	-	-
	(2.66)			(2.79)		
N	771	380	391	1170	589	581
Anderson LM χ^2 (1)	236.7	104.8	130.9	353.5	147.6	204.7
[*p*]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]

Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients. The specification also includes the number of adult men and women in the household age 15 and above and their average age. The identifying instrument is the average *As* concentration residual for genetically-linked individuals not residing in the same village as the subject.

Table 7
Estimates of the Effect of *As* Concentrations on the (Log) Male Daily Wage Rates: Men Aged 15-29

Measure	Worked for Wages	Log Wage	
Estimation procedure	ML Probit	IV	IV
As (ppb x 10^{-3})	-	-.0554 (2.07)	-.0523 (1.98)
As residual (ppb) in non-coresident lineage	.0731 (0.99)	-	-
Age	.599 (2.97)	.168 (2.47)	.254 (2.96)
Age squared	-.00928 (1.95)	-.00289 (1.91)	-.00445 (2.50)
Value of household landholdings (x 10^{-6})	-.411 (4.69)	-	-
Number of males in the household	-.0467 (0.54)	-	-
Average age of males	-.0181 (1.00)	-	-
Number of females in the household	-.105 (0.79)	-	-
Average age of females	-.00879 (0.45)	-	-
λ	-	-	.224 (1.64)
ρ	-	-	.495
N	589	425	425
Anderson LM χ^2 (1) [*p*]	-	113.9 [.000]	115.3 [.000]

Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients.

Table 8
FE (Village) Estimates of the Effects of Household Land Value and Gender on Individual Food Consumption, by Food Group

Food group/variable	Log Vegetables	Log Green Vegetables	Log Fruits	Log Grains	Log Tubers
Land value (x 10^{-6})	.171	-.109	.0406	-.0197	.0034
	(3.20)	(1.43)	(0.62)	(0.44)	(0.10)
Female	-.166	.00146	-.0232	-.185	-.161
	(4.67)	(0.04)	(0.69)	(16.2)	(7.29)
N	6,586	6,586	6,586	6,586	6,586
F(4, 1882)	192.5	38.0	2.67	863.3	356.9

Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients. The specification also includes age and age squared.

Appendix Table A
Estimates of the Determinants of *As* Concentrations (ppb):
Respondents Aged 15-29, by Gender

Group	All	Males	Females
As residual (ppb) in non-coresident lineage	1138 (22.4)	992.7 (13.9)	1262 (17.6)
Age	93.3 (0.73)	-112.8 (0.65)	263.7 (1.37)
Age squared	-1.79 (0.61)	2.51 (0.63)	-5.64 (1.27)
Value of household landholdings (x 10^{-6})	70.3 (1.13)	70.9 (0.98)	116.3 (1.01)
Number of males in the household	12.6 (0.19)	-36.2 (0.42)	130.2 (1.19)
Average age of males	22.6 (1.73)	5.69 (0.33)	42.6 (2.07)
Number of females in the household	-71.1 (0.80)	-97.7 (0.74)	-133.7 (0.93)
Average age of females	-2.23 (0.17)	.754 (0.04)	4.87 (0.25)
N	1170	589	581
F-statistic (9, N-10) [*p*]	58.3 [.000]	33.5 [.000]	56.6 [.000]

Absolute values of *t*-ratios in parentheses beneath coefficients.

Appendix Table B
Estimates of the Effects of *As* Concentrations on Cognition and Strength, by Gender and Estimation Procedure:
Men and Women Aged 15-29

Measure	Abridged Raven's CPM			Log Pinch Strength			Log Grip Strength		
Group/Est. Procedure	All	Males	Females	All	Males	Females	All	Males	Females
IV	-.155	-.259	-.0754	-.00845	-.0183	.00169	-.0121	-.0192	-.00369
	(2.64)	(2.55)	(1.08)	(1.73)	(2.20)	(0.29)	(1.63)	(1.69)	(0.39)
OLS	-.108	-.175	-.0479	-.00426	-.00717	.000914	-.00167	-.00771	.00658
	(3.38)	(3.50)	(1.14)	(1.58)	(1.73)	(0.26)	(0.41)	(1.34)	(1.15)

Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients. The specification also includes the number of adult men and women in the household over age 14 and their average age.

Appendix Table C
Estimates of the Effects of *As* Concentrations on Nutritional Status, by Gender and Estimation Procedure:
Men and Women Aged 15-29

Measure	Height			Weight			Weight/Height		
Group/Est. Procedure	All	Males	Females	All	Males	Females	All	Males	Females
IV	-.344	-.556	-.129	-.512	-.395	-.529	-.00263	-.00143	-.00321
	(1.91)	(1.74)	(0.64)	(2.61)	(1.20)	(2.22)	(2.39)	(0.81)	(2.25)
OLS	-.214	-.360	-.0400	-.403	-.425	-.330	-.00218	-.00203	-.00206
	(2.16)	(2.24)	(0.33)	(3.73)	(2.56)	(2.32)	(3.58)	(2.29)	(2.43)

Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients. The specification also includes the number of adult men and women in the household over age 14 and their average age

Appendix Table D
Estimates of the Effects of *As* Concentrations on Schooling for Men and Women and Wages for Men, by Estimation Procedure

Variable	Attending School			Years of Schooling Completed			Log Wage
Group/Est. Procedure	All	Males	Females	All	Males	Females	Males
IV	-.0337	-.0450	-.0274	-.0413	-.199	.0711	-.0554
	(2.24)	(1.76)	(1.56)	(0.44)	(1.23)	(0.63)	(2.07)
OLS	-.00486	.00277	-.0115	.00824	-.114	.116	-.0331
	(0.58)	(0.21)	(1.12)	(0.16)	(1.40)	(1.72)	(3.28)

Absolute values of asymptotic *t*-ratios in parentheses beneath coefficients. The specification also includes the number of adult men and women in the household age 15 and above and their average age.

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Subscribe to our newsletter
and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International
Growth Centre

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid
from the British people

Designed by soapbox.co.uk