

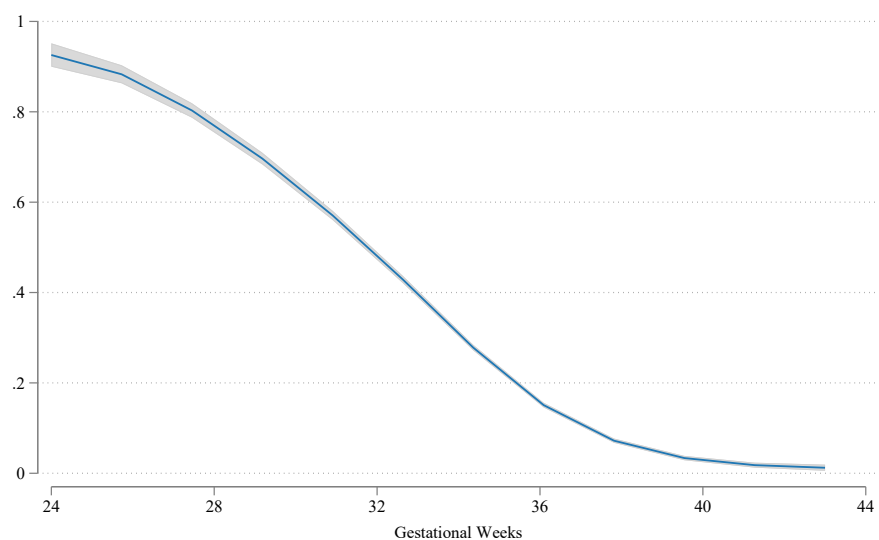
Supplemental Appendix

Luck of the Draw: The Causal Effect of Physicians on Birth Outcomes

Christian Posso, Jorge Tamayo, Arlen Guarín, Estefanía Saravia

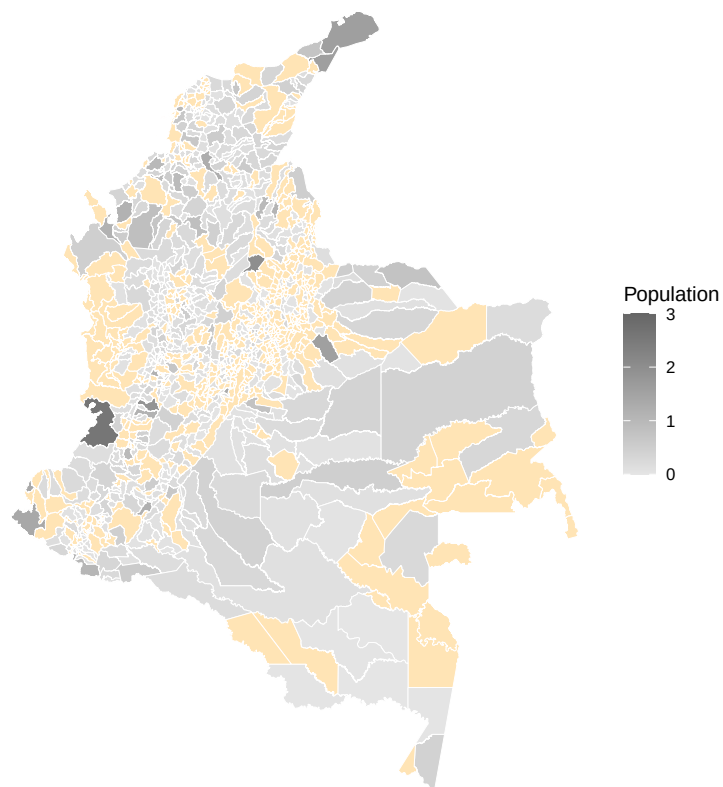
A Appendix

Figure A.1: Probability of low birth weight vs. gestational weeks, 2009-2012



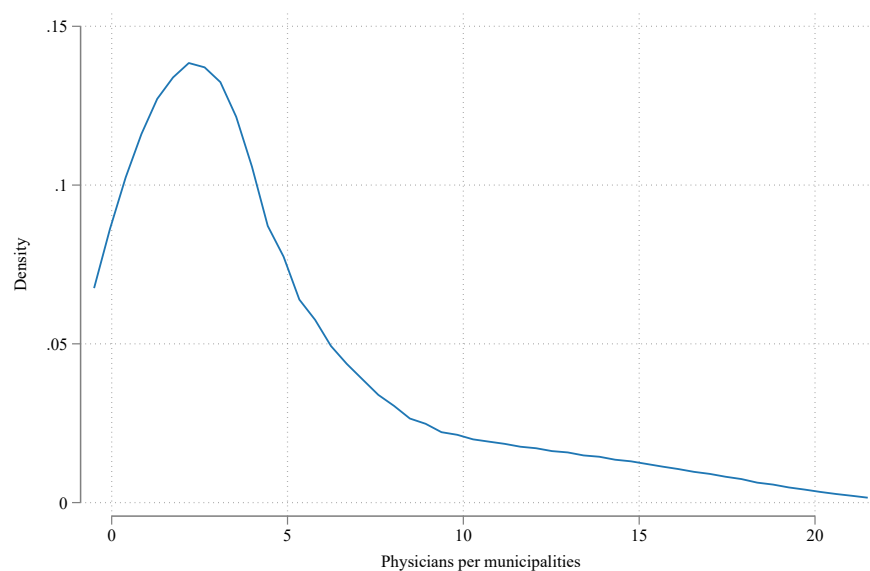
Notes: This figure presents the local polynomial regression fit of the probability of having low birth weight over the number of gestational weeks using all birth records for Colombia from 2009 to 2012.

Figure A.2: Population (per 100,000) for municipalities included in our main sample



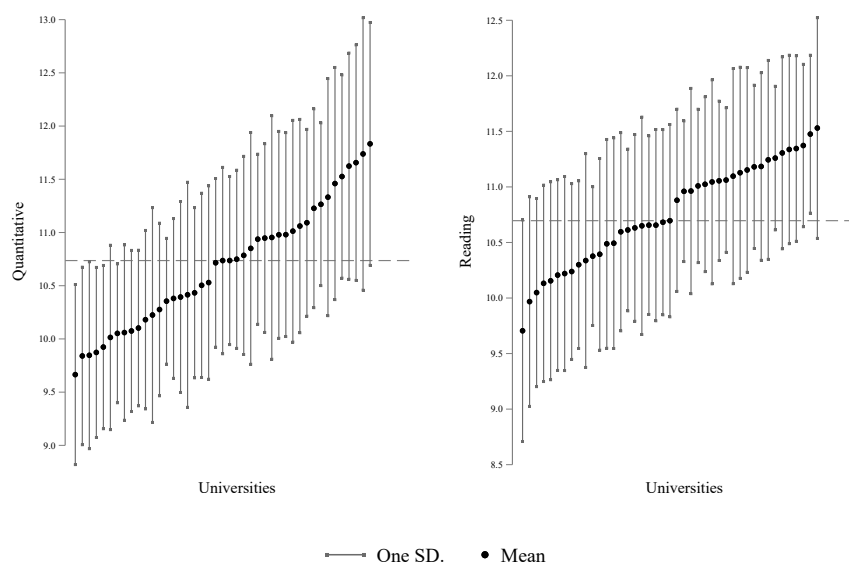
Notes: This figure presents the map ([Administrative Department of Statistics, 2018](#)) of the population per 100,000 people for the municipalities included in our main sample in 2005. The municipalities in orange are not included in our sample or do not have SSO.

Figure A.3: Distribution of physicians per municipality



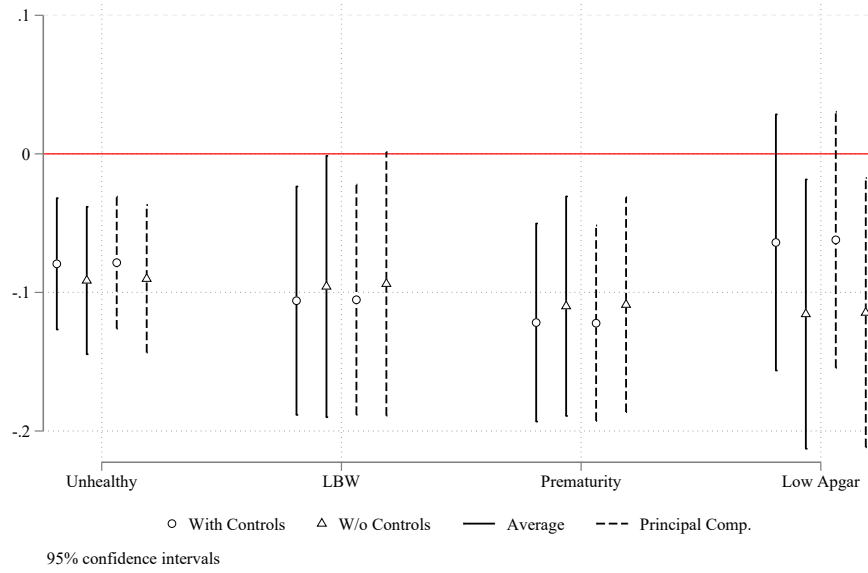
Notes: This figure shows the distribution of physicians per municipality for the sample of 582 municipalities with only one LHC. The data span from January 2012 to December 2012.

Figure A.4: Heterogeneity in quantitative and reading SABER PRO scores



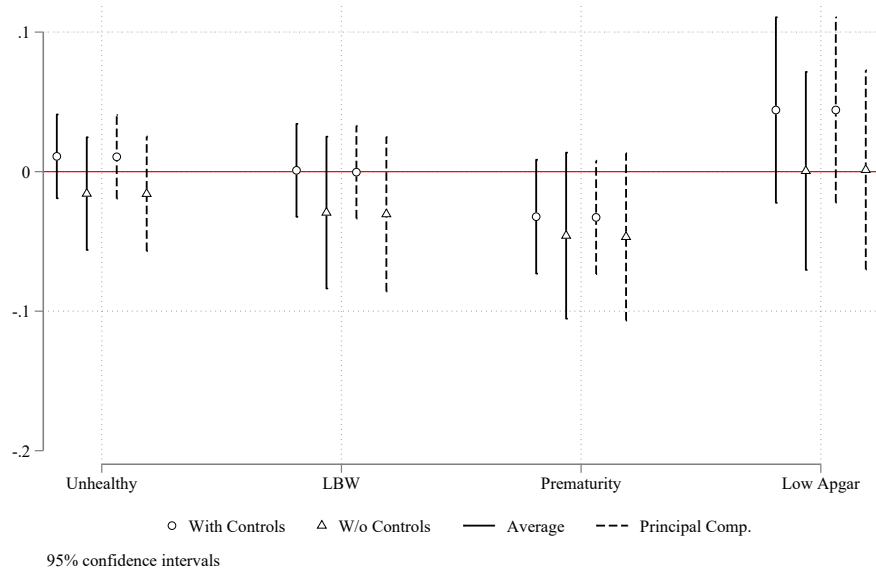
Notes: This figure reports the quantitative and reading test scores for the universities ([Ministry of Education, 2019](#)) that the physicians in our sample attended. The dataset covers 44 different universities. The figure shows the mean score for each university/program and an interval of one standard deviation to each side of the average. The dashed horizontal line represents the overall 50th percentile. The figure shows substantial heterogeneity both within and between programs. For all the fields reported, there is a difference of almost two standard deviations between the averages of the best and the worst programs.

Figure A.5: Main estimates using all sample



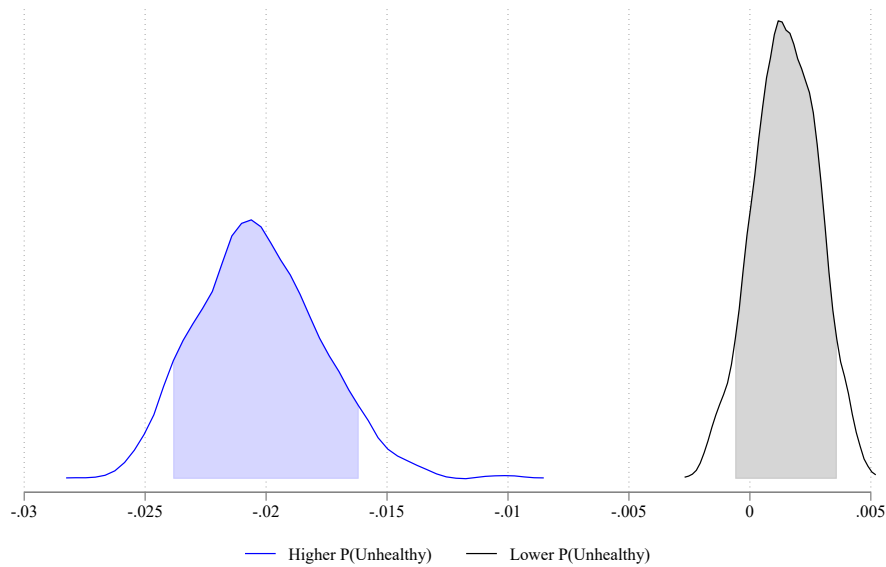
Notes: This figure presents our main estimates from equation (4). The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. First stage coefficient and standard error is shown in figure 2. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that takes the value of 1 if the LHC is above the 75th percentile of the distribution of low birth weight measured in 2010-2012 and zero otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010-2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010-2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010-2012, and 0 otherwise. These results show that the estimated effects are robust to the inclusion/exclusion of controls and the way we measure physicians' skills (Averages vs. principal components). Standard errors are clustered at the LHC level. 95% confidence intervals.

Figure A.6: Placebo using all samples and average scores



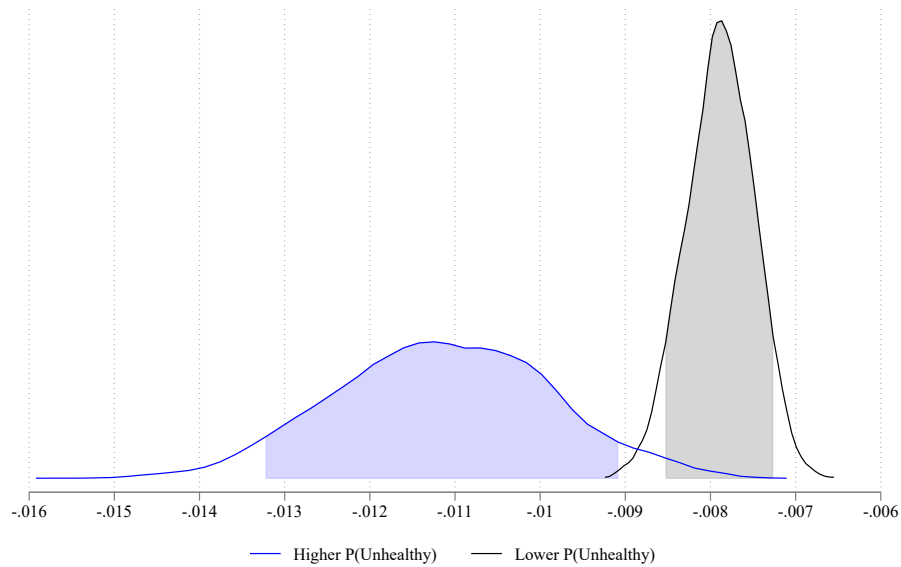
Notes: This figure presents our placebo estimates from equation (4). The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. First stage coefficient and standard error is shown in figure 2. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. These results show that the estimated effects are robust to the inclusion/exclusion of controls and the way we measure skills. These results support the ones presented in Table 3 on the robustness of the estimated zero effect for the placebo tests.

Figure A.7: Distribution of Logit simulations on antenatal consultations by predicted probability of unhealthy newborn



Notes: This figure plots the distribution of the estimated effects of physicians on antenatal consultations by mother's predicted probability of giving birth to an unhealthy child from 1,000 different random repetitions. In each of the 1,000 repetitions, to predict the probability of an unhealthy child, we divided our data into training and testing subsets of randomly selected LHCs using a K-means algorithm. On the training sets, we run a Logit model of the probability of being born unhealthy on our usual set of mother and LHC ex-ante covariates, and use the estimations to predict the probability of giving birth to an unhealthy child on each testing subset. Using the prediction on the testing sample, we divide each subset into high and low predicted probability of giving birth to an unhealthy child, defined as mothers with a probability of an unhealthy child below and above the 75th percentile, respectively. *Unhealthy* is a binary variable that takes the value of 1 if the newborn has low birth weight or if the newborn is premature (fewer than 37 weeks of gestation) or if the Apgar score of the newborn is lower than 7, and zero otherwise. The plotted coefficients represent the effect of being assigned a physician with one standard deviation higher quality (proxied by the average score) on the probability of having insufficient (less than four) antenatal consultations. All regressions control for draw-by-state fixed effects. The figure shows that there is almost no overlap between the distributions and that most of the mass of the distribution for the coefficient associated with the low predicted Unhealthy is around zero. This is consistent with the idea that more skilled physicians are better at targeting the care towards the more vulnerable mothers.

Figure A.8: Distribution of the coefficient logit simulations on the probability of being born unhealthy by the (ex-ante) predicted probability of an unhealthy newborn



Notes: This figure plots the distribution of the estimated effects of physicians on the probability of being born unhealthy by mother's predicted probability of giving birth to an unhealthy child from 1,000 different random repetitions. In each of the 1,000 repetitions, to predict the probability of an unhealthy child, we divided our data into training and testing subsets of randomly selected LHCs using a K-means algorithm. On the training sets, we run a Logit model of the probability of being born unhealthy on our usual set of mother and LHC ex-ante covariates, and use the estimations to predict the probability of giving birth to an unhealthy child on each testing subset. Using the prediction on the testing sample, we divide each subset into high and low predicted probability of giving birth to an unhealthy child, defined as mothers with a probability of an unhealthy child below and above the 75th percentile, respectively. *Unhealthy* is a binary variable that takes the value of 1 if the newborn has low birth weight or if the newborn is premature (fewer than 37 weeks of gestation) or if the Apgar score of the newborn is lower than 7, and zero otherwise. The plotted coefficients represent the effect of being assigned a physician with one standard deviation higher quality (proxied by the average score) on the probability of having insufficient (less than four) antenatal consultations. All regressions control for draw-by-state fixed effects. The figure shows that there is almost no overlap between the distributions and that the estimated effects of the more skilled physicians are consistently stronger for the population with higher predicted probability of being born unhealthy.

Table A.1: Summary statistics - physicians in the main sample

Covariate	Mean	Standard error
Sex (female)	0.558	0.497
The household has a private car	0.483	0.500
Number of people in the household	4.025	1.659
Father with tertiary education	0.644	0.479
Mother with tertiary education	0.634	0.482
Socioeconomic strata: 1 or 2 or rural areas	0.292	0.455
Socioeconomic strata: 4, 5 or 6	0.349	0.477
The household has internet	0.831	0.375
Monthly household income: Less than 2 MW	0.229	0.420
Monthly household income: between 2 and 3 MW	0.220	0.414
The father or the mother has a job	0.872	0.335
The household has a washing machine	0.854	0.353
The household has a television	0.859	0.348
The household has a cellphone	0.963	0.188
The house has proper flooring	0.908	0.289
The household has an oven	0.671	0.470
Physician's score on the Health care test	10.426	1.059
Physician's score on the Disease prevention test	10.431	1.010
Physician's score on the Reading test	10.624	1.007
Physician's score on the Math test	10.572	1.123
Physician's average score on SABER PRO	10.513	0.854
Observations	2,126	

Notes: This table reports the summary statistics for the physicians included in our main sample. These characteristics are obtained at the time physicians took their SABER PRO exam (before the SSO). *Sex* is a binary variable that takes the value of 1 if the physician is female and zero otherwise; *The household has a private car* that takes the value of 1 if the household of the physician owns a private car at the time the physician took the SABER PRO test and zero otherwise; *Number of people in the household* counts the number of individuals living in the same house as the physician; *Father with tertiary education* is a binary variable that takes the value of 1 if the physician's father has at least tertiary education and zero otherwise; *Mother with tertiary education* is a binary variable that takes the value of 1 if the physician's mother has at least tertiary education and zero otherwise; *Socioeconomic strata: 1 or 2 or rural areas* takes the value of 1 if the physician's household's socioeconomic strata at the time of the SABER PRO test was 1, 2 or rural and zero otherwise; *Socioeconomic strata: 4, 5 or 6* is a variable that takes the value of 1 if the physician's household's socioeconomic strata at the time of the SABER PRO test was 4, 5 or 6 and zero otherwise; *The household has internet* takes the value of 1 if the physician had internet service at home at the time of the test; *Monthly household income: Less than 2MW* takes the value of 1 if the physician's household had an income lower than 2 times the national monthly minimum wage and zero otherwise; *Monthly household income: between 2 and 3 MW* takes the value of 1 if the physician's household had an income between 2 and 3 times the national monthly minimum wage and zero otherwise; *The father or the mother has a job* takes value 1 if either of the physician's parents have a job; The household has a washing machine, television, cellphone, proper flooring or oven, take the value of 1 if the household has that characteristic described and zero otherwise; physician's scores are continuous variables of the score obtained on each SABER PRO test subgroup; physician's average score on SABER PRO is the average of the four main components of the test, health care, disease prevention, reading, and math.

Table A.2: Covariate balance at LHC level using all the areas tested in the SABER PRO

Covariate	Coefficient	Standard Error
<i>a. Pretreatment variables (2010-2012)</i>		
Unhealthy	0.0060	0.0094
Low birth weight	−0.0027	0.0026
Prematurity	0.0044	0.0043
Low Apgar score	−0.0020	0.0048
Insufficient prenatal care (Prop.)	−0.0054	0.0078
Female infants	−0.0011	0.0046
Mothers with basic education	−0.0035	0.0082
Married mothers	−0.0040	0.0070
Teenage mothers	0.0055	0.0053
Number of LHCs per municipality	−0.0247	0.0203
Municipality population	1,196.04	2,767.35
<i>b. Concurrent variables</i>		
Female newborn	−0.0008	0.0033
Mother with basic education	0.0053	0.0077
Married mother	−0.0034	0.0059
Teenage mother	0.0006	0.0039
Number of LHCs per municipality	−0.0043	0.0209
Municipality population	1,260.14	2,807.26

Notes: This table presents the results of different LHC-by-cohort level regressions (equation 5) of the LHC-level variables, listed in the first column, on the measure of physicians' skill level and the draw-by-state fixed effects. The coefficient and the standard error of the physicians' skill variable are reported in the second and third columns, respectively. Standard errors are clustered at the LHC level. LHCs' characteristics in panel a come from the 2010–2012 DANE VSRs, using a total of 1,837 LHC-by-cohort observations. LHCs' characteristics in panel b come from the 2013–2015 DANE VSRs, using a total of 1,714 LHC-by-cohort observations. *Unhealthy*, our main measure of health at birth, is the proportion of newborn infants with at least one of the three following conditions: low birth weight, prematurity, or low Apgar score. *Low birth weight* is the proportion of newborn infants whose birth weight was less than 2,500 grams. *Prematurity* is the proportion of newborn infants who were born after fewer than 37 weeks of gestation. *Low Apgar score* is the proportion of newborn infants whose Apgar score was lower than 7. *Insufficient prenatal care* is the proportion of mothers who had fewer than four prenatal checkups. *Female infants* is the proportion of female infants. *Mothers with basic education* is the proportion of mothers with at least secondary education at the time they gave birth. *Married mothers* is the proportion of mothers who were married at the time they gave birth. *Teenage mothers* is the proportion of mothers who were 19 years old or younger at the time they gave birth. *Number of LHCs per municipality* is the count of LHCs in the birthplace municipality. We interpret the non-significance of these estimates as evidence in favor of the randomness of the assignment of physicians.

Table A.3: Placebo other years

	Unhealthy		LBW		Prematurity		Apgar < 7	
	Average Health Scores (1)	PCA Health Scores (2)	Average Health Scores (3)	PCA Health Scores (4)	Average Health Scores (5)	PCA Health Scores (6)	Average Health Scores (7)	PCA Health Scores (8)
<i>a. 2 years</i>								
Coefficient	−0.0029	−0.0027	−0.003*	−0.0029*	−0.0005	−0.0004	−0.0022	−0.0021
SE	(0.0026)	(0.0026)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0019)	(0.0019)
Relative effect	−2.90%	−2.72%	−6.57%	−6.39%	−1.09%	−0.98%	−5.44%	−5.23%
<i>b. 2.5 years</i>								
Coefficient	−0.0022	−0.0021	−0.0009	−0.0008	−0.0006	−0.0006	−0.0022	−0.0022
SE	(0.0023)	(0.0023)	(0.0013)	(0.0013)	(0.0015)	(0.0015)	(0.0019)	(0.0019)
Relative effect	−2.18%	−2.07%	−1.95%	−1.82%	−1.41%	−1.35%	−5.58%	−5.42%
<i>c. 3 years</i>								
Coefficient	−0.0035	−0.0034	−0.0007	−0.0007	−0.0016	−0.0016	−0.0026	−0.0026
SE	(0.0022)	(0.0022)	(0.0012)	(0.0013)	(0.0014)	(0.0014)	(0.0018)	(0.0018)
Relative effect	−3.33%	−3.30%	−1.40%	−1.45%	−3.85%	−3.92%	−6.27%	−6.12%
<i>d. 3.5 years</i>								
Coefficient	−0.0014	−0.0014	−0.0007	−0.0007	−0.0012	−0.0013	−0.0006	−0.0006
SE	(0.0023)	(0.0023)	(0.0011)	(0.0012)	(0.0016)	(0.0016)	(0.0018)	(0.0018)
Relative effect	−1.26%	−1.29%	−1.41%	−1.58%	−2.80%	−2.87%	−1.47%	−1.37%

Notes: This table presents placebo test in which we estimate equation (4) but moving the arrival date 3.5, 3, 2.5 and 2 years before the start of the SSO program. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. First stage coefficient and standard error is shown in figure 2. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. The numbers in parentheses are LHC-level-clustered standard errors. The results show that regardless of the time window that we use for the calculation of the placebo test, the estimated coefficients are always precisely estimated as zeros, which we interpret as evidence of the randomness of the assignment of the physicians to the LHCs.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Placebo robustness checks

	Unhealthy		LBW		Prematurity		Apgar < 7	
	Average Health Scores (1)	PCA Health Scores (2)	Average Health Scores (3)	PCA Health Scores (4)	Average Health Scores (5)	PCA Health Scores (6)	Average Health Scores (7)	PCA Health Scores (8)
a. Without controls								
Coefficient	−0.0019	−0.0019	−0.0014	−0.0014	−0.0024	−0.0024	0.0001	0.0001
SE	(0.0024)	(0.0025)	(0.0013)	(0.0013)	(0.0016)	(0.0016)	(0.0017)	(0.0017)
Relative effect	−1.58%	−1.59%	−2.93%	−3.05%	−4.59%	−4.68%	0.05%	0.13%
b. With controls								
Coefficient	0.0013	0.0013	0.0001	−0.0001	−0.0017	−0.0017	0.0020	0.0020
SE	(0.0018)	(0.0018)	(0.0008)	(0.0008)	(0.0011)	(0.0011)	(0.0016)	(0.0016)
Relative effect	1.09%	1.06%	0.09%	−0.04%	−3.23%	−3.28%	4.42%	4.42%
Average dependent variable	0.118		0.046		0.052		0.046	
Number of observations	261,616							

Notes: This table presents our placebo estimates from equation (4) with and without controls. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. First stage coefficient and standard error are shown in Figure 2. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. Note that the results are robust to the inclusion/exclusion of controls and how we measure skills. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Main estimates using all the areas tested in the SABER PRO

Dependent variable: Unhealthy							
	Average All	Average Health Scores	Health Care Score	Disease Prevention Score	Average Academic Scores	Reading Score	Quantitative Score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>a. Without controls</i>							
Coefficient	-0.0109***	-0.0087***	-0.0089***	-0.0054*	-0.0105***	-0.0065**	-0.0106***
SE	(0.0026)	(0.0026)	(0.0026)	(0.0027)	(0.0027)	(0.0027)	(0.0024)
Relative effect	-11.44%	-9.14%	-9.33%	-5.64%	-11.01%	-6.78%	-11.12%
<i>b. With controls</i>							
Coefficient	-0.0096***	-0.0076***	-0.0075***	-0.0051**	-0.0094***	-0.0066**	-0.0091***
SE	(0.0022)	(0.0023)	(0.0023)	(0.0025)	(0.0023)	(0.0026)	(0.0022)
Relative effect	-10.11%	-7.94%	-7.90%	-5.33%	-9.86%	-6.89%	-9.56%
Average dependent variable				0.095			
Number of observations				255,089			

Notes: This table presents our main estimates from equation (4) using all areas tested in the SABER PRO. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status; number of inhabitants in the municipality; number of LHCs per municipality; area; an indicator variable that takes the value of 1 if the LHC is above the 75th percentile of the distribution of low birth weight measured in 2010-2012 and zero otherwise; an indicator variable that takes the value of 1 if the LHC is above the 75th percentile of the distribution of prematurity measured in 2010-2012 and zero otherwise; and an indicator variable that takes the value of 1 if the LHC is above the 75th percentile of the distribution of the Apgar score measured in 2010-2012 and zero otherwise. These results show that the estimated effects are robust to using the average of the four areas tested in the SABER PRO (health management, public health, reading, quantitative) as well as each individual (except for reading) score as proxies of the physician's skills before the SSO program. The results are also robust to the inclusion/exclusion of controls and how we measure skills. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Cohort-level mortality estimates

	Fetal deaths		Fetal and neonatal deaths		Infant Mortality Ratio	
	Average Health Scores	PCA Health Scores	Average Health Scores	PCA Health Scores	Average Health Scores	PCA Health Scores
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	-0.9408	-0.6315	-0.9582	-0.6421	-0.0005	-0.0003
SE	(2.4024)	(1.5996)	(2.4228)	(1.613)	(0.0015)	(0.0010)
Relative effect	-6.39%	-4.29%	-6.10%	-4.09%	-2.18%	-1.43%
Average dependent variable	14.728		15.705		0.024	
Number of observations	1,073		1,073		1,073	

Notes: This table presents our cohort-level estimates on mortality following equation (4). Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. *Fetal deaths* is the total number of fetal deaths registered at the LHC during the timeframe when the cohort was assigned. *Fetal and neonatal deaths* represent the total number of fetal deaths and fatalities of children under one year old registered in an LHC during the cohort's assignment period (ideally, we would have preferred to focus on shorter-term mortality, but under one year was the most granular definition of infant mortality available in our data). *Infant Mortality Ratio*, represents the number of fetal and neonatal deaths divided by the total number of births during the cohort's assignment period. These variables are regressed on either the cohort's average health score (columns 1, 3, 5) or the cohort's PCA for the health scores (columns 2, 4, 6). We restrict to cohorts assigned to LHC where there are at least 5 births during their assignment period, but the results are similar when this threshold is increased/decreased/ignored. While these results are expected to be subject to high measurement error attenuation bias, we still observe negative, albeit not statistically significant, point estimates, which aligns with our main results.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Controlling by Share of SSOs on LHC

	Unhealthy	LBW	Prematurity	Apgar < 7
	(1)	(2)	(3)	(4)
Coefficient	-0.0087***	-0.0041**	-0.0045***	-0.0044**
SE	(0.0026)	(0.0020)	(0.0016)	(0.0019)
Relative effect	-9.18%	-9.52%	-10.91%	-11.77%
Average dependent variable	0.095	0.043	0.041	0.038
Number of observations		255,089		

Notes: This table presents our main estimates from equation (4) controlling for the share of SSO physicians at the LHC. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. First stage coefficient and standard error are shown in figure 2. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Numbers in parentheses are LHC-level clustered standard errors. We interpret the high significance and consistency of these results across the different measures of health at birth as evidence of the important role that skilled physicians play in determining an infant's health at birth.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Antenatal consultations < 4

Dependent variable: <i>Antenatal consultations < 4</i>		
	Average Health Scores (1)	PCA Health Scores (2)
<i>a. Without controls</i>		
Coefficient	-0.0029	-0.0031
SE	(0.0093)	(0.0094)
Relative effect	-1.77%	-1.93%
<i>b. With controls</i>		
Coefficient	-0.0055	-0.0058
SE	(0.0092)	(0.0093)
Relative effect	-3.39%	-3.53%
Average dependent variable	0.163	
Number of observations	255,089	

Notes: This figure presents our main estimates from equation (4). The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation, on the probability that mothers are scheduled for less than four prenatal checkups (Insufficient antenatal consultations). Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. First stage coefficient and standard error are shown in figure 2. *Antenatal consultations < 4* takes the value one if the mother attended fewer than 4 consultations while pregnant, and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. Note that the results are robust to the inclusion/exclusion of controls and how we measure skills. Numbers in parentheses are LHC-level clustered standard errors. The results show no significant average effect of more-skilled doctors on the probability that mothers are scheduled for less than four prenatal checkups.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: Main estimates without and with controls

	Unhealthy		LBW		Prematurity		Apgar < 7	
	Average Health Scores (1)	PCA Health Scores (2)	Average Health Scores (3)	PCA Health Scores (4)	Average Health Scores (5)	PCA Health Scores (6)	Average Health Scores (7)	PCA Health Scores (8)
a. Without controls								
Coefficient	-0.0087***	-0.0086***	-0.0041**	-0.0040*	-0.0045***	-0.0045***	-0.0043**	-0.0043**
SE	(0.0026)	(0.0026)	(0.0021)	(0.0021)	(0.0017)	(0.0016)	(0.0019)	(0.0019)
Relative effect	-9.14%	-9.02%	-9.57%	-9.38%	-10.99%	-10.89%	-11.56%	-11.46%
b. With controls								
Coefficient	-0.0076***	-0.0075***	-0.0045**	-0.0045**	-0.0050***	-0.0050***	-0.0024	-0.0023
SE	(0.0023)	(0.0023)	(0.0018)	(0.0018)	(0.0015)	(0.0015)	(0.0018)	(0.0018)
Relative effect	-7.94%	-7.85%	-10.60%	-10.53%	-12.17%	-12.22%	-6.39%	-6.20%
Average dependent variable	0.095		0.043		0.041		0.038	
Number of observations	255,089							

Notes: This table presents our main estimates from equation (4) with and without controls. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. First stage coefficient and standard error are shown in Figure 2. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. These results show that the estimated effects are robust to the inclusion/exclusion of controls and the way we measure quality. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: Main results using covariance index ([Anderson, 2008](#))

	Unhealthy Cov index		Unhealthy standardized	
	Average Scores (1)	PCA Scores (2)	Average Scores (3)	PCA Scores (4)
<i>a. Without controls</i>				
Coefficient	-0.0220***	-0.0218***	-0.0297***	-0.0292***
SE	(0.0065)	(0.0065)	(0.0088)	(0.0088)
Relative effect	-3.23%	-3.19%	-2.97%	-2.92%
<i>b. With controls</i>				
Coefficient	-0.0192***	-0.0190***	-0.0257***	-0.0255***
SE	(0.0060)	(0.0060)	(0.0078)	(0.0079)
Relative effect	-2.82%	-2.79%	-2.57%	-2.55%
Number of observations	255,089			

Notes: This table presents our main estimates from equation (4) using [Anderson \(2008\)](#) covariance index. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. These results show that the estimated effects are robust to using the covariance index as an outcome instead of unhealthy. The results are also robust to the inclusion/exclusion of controls and how we measure quality. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.11: Main estimates using a Logit model

	Unhealthy		LBW		Prematurity		Apgar < 7	
	Average Health Scores (1)	PCA Health Scores (2)	Average Health Scores (3)	PCA Health Scores (4)	Average Health Scores (5)	PCA Health Scores (6)	Average Health Scores (7)	PCA Health Scores (8)
a. Without controls								
Coefficient	-0.0070***	-0.0069***	-0.0032**	-0.0031**	-0.0039***	-0.0039***	-0.0034**	-0.0033**
SE	(0.0021)	(0.0021)	(0.0015)	(0.0015)	(0.0014)	(0.0014)	(0.0015)	(0.0015)
Relative effect	-7.34%	-7.22%	-7.45%	-7.27%	-9.54%	-9.41%	-9.03%	-8.92%
b. With controls								
Coefficient	-0.0059***	-0.0058***	-0.0036***	-0.0035***	-0.0046***	-0.0046***	-0.0020	-0.0019
SE	(0.0018)	(0.0018)	(0.0013)	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Relative effect	-6.20%	-6.13%	-8.37%	-8.29%	-11.17%	-11.18%	-5.22%	-5.08%
Average dependent variable	0.095		0.043		0.041		0.038	
Number of observations	255,079							

Notes: This table presents our main estimates from equation (3) using a logit model. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. These results show that the estimated effects are robust to using an analogous Logit model and computing the average marginal effect associated with an increase in one standard deviation of the skill measure. The results are also robust to the inclusion/exclusion of controls and how we measure quality. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.12: Main estimates linearity

		Dependent variable: <i>Unhealthy</i>	
		Average Health Scores (1)	PCA Health Scores (2)
Quartile 2	Coefficient	-0.0150	-0.0147
	SE	(0.010)	(0.0105)
	Relative effect	-15.79%	-15.45%
Quartile 3	Coefficient	-0.0217**	-0.0227
	SE	(0.0104)	(0.0145)
	Relative effect	-22.74%	-23.82%
Quartile 4	Coefficient	-0.0196**	-0.0199**
	SE	(0.0089)	(0.0079)
	Relative effect	-20.62%	-20.94%

Notes: This table presents our main estimates from equation (4) using the quartiles of the quality distribution. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level was at the 2nd, 3rd, or 4th quartile of the physicians' quality distribution compared to being treated at an LHC that was randomly assigned SSO physicians whose skill level was at the 1st quartile. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Numbers in parentheses are LHC-level clustered standard errors. While not all the coefficients are statistically different from each other, we do observe increases in the point estimates associated with higher quartiles and cannot discard the linearity of the effects.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.13: Main estimates without, with dummy and continuous controls

	Unhealthy		LBW		Prematurity		Apgar < 7	
	Average Health Scores (1)	PCA Health Scores (2)	Average Health Scores (3)	PCA Health Scores (4)	Average Health Scores (5)	PCA Health Scores (6)	Average Health Scores (7)	PCA Health Scores (8)
a. Without controls								
Coefficient	-0.0087***	-0.0086***	-0.0041**	-0.0040*	-0.0045***	-0.0045***	-0.0043**	-0.0043**
SE	(0.0026)	(0.0026)	(0.0021)	(0.0021)	(0.0017)	(0.0016)	(0.0019)	(0.0019)
Relative effect	-9.14%	-9.02%	-9.57%	-9.38%	-10.99%	-10.89%	-11.56%	-11.46%
b. With dummy controls								
Coefficient	-0.0076***	-0.0075***	-0.0045**	-0.0045**	-0.0050***	-0.0050***	-0.0024	-0.0023
SE	(0.0023)	(0.0023)	(0.0018)	(0.0018)	(0.0015)	(0.0015)	(0.0018)	(0.0018)
Relative effect	-7.94%	-7.85%	-10.60%	-10.53%	-12.17%	-12.22%	-6.39%	-6.20%
c. With continuous controls								
Coefficient	-0.0077***	-0.0077***	-0.0036**	-0.0036**	-0.0041***	-0.0042***	-0.0037**	-0.0037**
SE	(0.0021)	(0.0021)	(0.0017)	(0.0017)	(0.0012)	(0.0012)	(0.0017)	(0.0017)
Relative effect	-8.11%	-8.06%	-8.53%	-8.52%	-9.95%	-10.12%	-9.95%	-9.73%
Average dependent variable	0.095		0.043		0.041		0.038	
Number of observations	255,089							

Notes: This table presents our main estimates from equation (4) without controls, and with dummy and continuous controls. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. The first stage coefficient and standard error are shown in Figure 2. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With dummy controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. Regressions for the coefficients labeled as *With continuous controls* include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is adolescent and zero otherwise; marital status; the LHC's low birth weight average measured in 2010–2012; the LHC's prematurity percentage measured in 2010–2012; and the LHC's Apgar average measured in 2010–2012 and zero otherwise. These results show that the estimated effects are robust to the inclusion/exclusion of controls and the way we measure skills. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.14: Interaction between cohort scores and program scores

		Unhealthy (1)	LBW (2)	Prematurity (3)	Apgar < 7 (4)
Average Health Score	Coefficient	-0.0090**	-0.0046**	-0.0047**	-0.0035
	SE	(0.0035)	(0.0019)	(0.0021)	(0.0029)
	Relative effect	-9.40%	-10.83%	-11.47%	-9.36%
Program Average	Coefficient	0.0023	0.0028**	0.0023	-0.0011
	SE	(0.0028)	(0.0014)	(0.0016)	(0.0026)
	Relative effect	2.39%	6.45%	5.53%	-2.83%
Av. Health Sc. x Prog. Av.	Coefficient	0.0013	0.0012	0.0014	0.0005
	SE	(0.0017)	(0.0015)	(0.0011)	(0.0013)
	Relative effect	1.35%	2.75%	3.32%	1.42%
Average dependent variable		0.095	0.043	0.041	0.038
Number of observations			255,089		

Notes: This table presents our main estimates from equation (4) using the interaction between cohort and program scores. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. These results show that the effects presented in Table 4 are not driven by top-ranked universities. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.15: Main results using municipalities with one LHC

	Unhealthy		LBW		Prematurity		Apgar < 7	
	Average Health Scores (1)	PCA Health Scores (2)	Average Health Scores (3)	PCA Health Scores (4)	Average Health Scores (5)	PCA Health Scores (6)	Average Health Scores (7)	PCA Health Scores (8)
a. Without controls								
Coefficient	-0.0077***	-0.0075***	-0.0040*	-0.0039*	-0.0045***	-0.0044***	-0.0033*	-0.0033*
SE	(0.0025)	(0.0026)	(0.0021)	(0.0021)	(0.0017)	(0.0017)	(0.0017)	(0.0017)
Relative effect	-7.97%	-7.84%	-9.21%	-8.95%	-10.69%	-10.57%	-8.71%	-8.72%
b. With controls								
Coefficient	-0.0068***	-0.0068***	-0.0038**	-0.0038**	-0.0046***	-0.0046***	-0.0023	-0.0023
SE	(0.0024)	(0.0024)	(0.0019)	(0.0019)	(0.0015)	(0.0015)	(0.0018)	(0.0018)
Relative effect	-7.07%	-7.02%	-8.91%	-8.88%	-10.89%	-11.02%	-6.14%	-6.01%
Average dependent variable	0.096		0.043		0.042		0.038	
Number of observations	238,296							

Notes: This table presents our main estimates from equation (4) using municipalities with only one LHC. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. The table shows that the results presented in Table 4 are almost identical if we exclude from our main sample the ten municipalities with more than two LHCs per municipality. The results are also robust to the inclusion/exclusion of controls and how we measure skills. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.16: Main results using the weighted score without and with controls

	Unhealthy	LBW	Prematurity	Apgar < 7
	(1)	(2)	(3)	(4)
Average Health Scores				
<i>a. Without controls</i>				
Coefficient	-0.0084***	-0.0041*	-0.0047***	-0.0041**
SE	(0.0026)	(0.0021)	(0.0017)	(0.0019)
Relative effect	-8.82%	-9.54%	-11.35%	-10.92%
<i>b. With controls</i>				
Coefficient	-0.0070***	-0.0044**	-0.0049***	-0.0021
SE	(0.0023)	(0.0018)	(0.0015)	(0.0018)
Relative effect	-7.39%	-10.18%	-12.00%	-5.70%
Average dependent variable	0.095	0.043	0.041	0.037
Number of observations		252,159		

Notes: This table presents our main estimates from equation (4) using the weighted score. The coefficients represent the effect of being treated at an LHC that was randomly assigned SSO physicians whose skill level is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. *Unhealthy* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams, if the newborn infant is born after fewer than 37 weeks of gestation, or if the Apgar score of the newborn infant is lower than 7 and zero otherwise. *LBW* is a binary variable that takes a value of 1 if the newborn infant has a birth weight below 2,500 grams and zero otherwise. *Prematurity* is a binary variable that takes a value of 1 if the newborn infant is born after fewer than 37 weeks of gestation and zero otherwise. *Low Apgar* is a binary variable that takes a value of 1 if the Apgar score of the newborn infant is lower than 7 and zero otherwise. All regressions control for draw-by-state fixed effects. Regressions for the coefficients labeled as *With controls* also include the following controls: an indicator variable for the sex of the newborn; an indicator variable that takes the value of 1 if the mother has at least secondary education and zero otherwise; an indicator variable that takes the value of 1 if the mother is 19 years old or younger and zero otherwise; marital status, number of inhabitants in the municipality; number of LHCs per municipality; an indicator variable that equals 1 if the LHC is above the 75th percentile of the low birth weight distribution for the country in 2010–2012, and 0 otherwise; an indicator variable that equals 1 if the LHC is above the 75th percentile of the prematurity distribution for the country in 2010–2012, and 0 otherwise; and an indicator variable that equals 1 if the LHC is above the 75th percentile of the Apgar score distribution for the country in 2010–2012, and 0 otherwise. The table shows that the results are very similar when the weighted score is used as a proxy for physicians' skills. The results are also robust to the inclusion/exclusion of controls and how we measure skills. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.17: Heterogeneity of the Effects on Birth Outcomes by SSO Share Quartile

	Q1	Q2	Q3	Q4
	Average Health Scores			
	(1)	(2)	(3)	(4)
Coefficient	−0.0102***	−0.0050	−0.0093	−0.0152**
SE	(0.0030)	(0.0062)	(0.0084)	(0.0064)
Relative effect	−10.82%	−5.24%	−9.48%	−15.14%
Average dependent variable	0.094	0.096	0.099	0.100
Number of observations	161,424	44,992	26,237	22,428

Note: This table presents heterogeneity estimates from equation (4) by quartiles of SSO physician share at the LHC. The coefficients represent the effect of being treated at an LHC in a given quartile of SSO share when the average skill level of SSO physicians is higher by one standard deviation. Relative (percent) effects are computed as the coefficient divided by the average of the dependent variable. First stage coefficient and standard error are shown in figure 2. All regressions control for draw-by-state fixed effects. Numbers in parentheses are LHC-level clustered standard errors.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$