

The Long-Term Impact of Work-Hour Regulations on Physician Labor Supply

Judith Liu

Department of Economics

Syracuse University

September 2018

Abstract

This paper estimates the impact of a transitory reduction in hours during physicians' early career on their long-term labor supply. I exploit the work-hour regulations that limit the maximum workweek by residents as the source of exogenous variation. The results show that exposure to the regulations significantly decreases practicing physicians' labor supply by about four hours per week on average, with female physicians being more responsive to a given reduction in early career hours. Distributional results using a changes-in-changes model confirm that the regulations primarily affect the upper end of the work hours distribution. To reveal potential mechanisms of these effects, I find that the reform increases the probabilities of marriage and having a child, as well as the total number of children, for female physicians. In contrast, it does not have a significant impact on marriage or fertility outcomes for male physicians. These findings provide a better understanding of physicians' hours of work in response to the reform over time and the role of gender with respect to labor supply behavior and fertility decisions.

Keywords: regulation, physician labor supply, gender differences, marriage, fertility

JEL classification: J12, J13, J16, J44

1. Introduction

The recent expansion of the health care system and forecasts of physician shortage have made the issue of physician labor supply increasingly important. Over the last three decades, average hours worked by physicians have been falling in many developed countries, including the United States, Canada, and Australia, among others (Buske, 2004; Scott, 2006; Watson et al., 2006; Crossley et al., 2009; Staiger et al., 2010). There has also been a dramatic change in the composition of the physician workforce, with the female share of medical students rising from 25 percent in 1976 to 48 percent in 2006 (Chen and Chevalier, 2012). Most research and policy debates have focused on physician supply at the extensive margin (the number of practicing physicians), whereas physician supply at the intensive margin (the amount of patient care hours or services provided by practicing physicians) has been understudied (Staiger et al., 2010). A better understanding of physicians' hours of work decisions and gender differences is crucial for human resource planning purposes in the health care sector.

An important determinant of individuals' hours of work is their work experience in early career, but there is little evidence on its long-term consequences. Empirically, it is particularly difficult to identify sources of exogenous variation to test for causal effects of a transient change in labor supply. The work-hour regulations that limit the maximum hours worked by residents provide a plausibly exogenous shift in physicians' early career hours.

To become a practicing physician in the U.S., an individual must complete three to

seven years of residency training after college and medical school, and then obtain medical licensure to practice medicine. Traditionally, long hours are a component of residency training, yet they may contribute to sleep deprivation which compromises patient safety. In order to reduce potential harm due to overwork of residents, the Accreditation Council for Graduate Medical Education (ACGME) imposed regulations that restrict the average hours worked by residents to 80 hours per week and enforce standards for their duty hours in 2003. A large number of studies have examined the effects of the ACGME regulations on patients' safety and health outcomes, as well as residents' education and well-being (Philibert et al., 2013); yet the impacts on physicians' labor market outcomes have not been thoroughly explored in the literature. To my knowledge, the only paper that uses the ACGME regulations as a natural experiment to estimate the effects of early career hours is Wasserman (2018), which focuses on changes in specialty choice across gender.

This paper investigates whether the reform affects physicians' hours of work after they complete residency and do not face the hours constraints anymore. Using data from the 1989-2017 monthly Current Population Surveys (CPS), my primary empirical strategy exploits the cohort-time variation in exposure to the ACGME regulations. As a result of the reform, the mean resident hours per week decrease by 10.54 for males and 5.94 for females. Using a difference-in-differences approach, the estimates suggest that exposure to the reform during residency significantly decreases mean hours worked after residency by about four hours per

week, and the effects are not statistically different between male and female physicians. When taking the effects of the reform on resident hours by gender into account, a given reduction in hours during residency decreases post-residency hours substantially more for females than for males.

Since the policy limits the maximum workweek by residents, it should primarily affect those who would have worked more than 80 hours per week during residency in the absence of the regulations. To account for this disproportionate impact at the upper end of the hours distribution, I use a changes-in-changes (CIC) approach proposed by Melly and Santangelo (2015), which estimates unconditional treatment effects of the whole distribution in the presence of covariates. Overall, the CIC estimates provide further evidence of the negative effects of the reform on long-term labor supply and show that such negative effects become stronger when moving towards the upper tail of the distribution for both male and female physicians. The greater impact among those with the highest hours confirms that the reform primarily affects the upper end of the hours distribution.

As well-documented in the literature related to physician labor supply, gender differences in hours of work may be attributable to child-rearing (e.g., Sasser, 2005; Wang and Sweetman, 2013; Goldin, 2014; Wasserman, 2018), which suggests a potential mechanism of the above long-term effects. To guide our understanding about the presence of this mechanism in this context, I further examine how the regulations affect male and female physicians'

marriage and fertility decisions. The results show that the reform increases the probabilities of marriage and having a child, as well as the total number of children, for female physicians. In contrast, the regulations do not have a significant impact on marriage and fertility outcomes for male physicians. These findings are consistent with previous studies and provide strong evidence on gender differences in fertility choices in response to a policy that reduces time requirements during the prime childbearing years.

A potential mechanism for the negative impact of the reform on male physicians' long-term labor supply is through human capital accumulation. The literature finds that the work-hour limits reduce continuity of care and educational continuity for residents in surgical specialties, which may lead to negative consequences for residents' professional development and preparedness for practice (Feanny et al., 2005; Vidyarthi et al., 2006; McBurney et al., 2008; Nakayama et al., 2009; Philibert et al., 2013). Since there are substantially more male physicians in surgical specialties, they experience a greater reduction in human capital accumulation, which in turn decreases the return to work and hours of work in the future. The negative impact of the reform on long-term labor supply for male physicians may be largely explained by this potential mechanism.

This paper has three main contributions to the existing literature. First, it exploits the work-hour regulations on residents to identify the effects of a reduction in early career hours on long-term labor supply. The findings aid our understanding of physicians' hours of work

decisions in response to the reform over time. Second, this paper explores gender differences in labor supply behaviors, along with marriage and fertility decisions as potential mechanisms. With the composition of the physician workforce changing dramatically, especially the increasing participation of females, a better understanding of how males and females respond to policy changes can suggest ways to orient policies more effectively. Third, this paper contributes to the research on dynamics of labor supply, for which it is often difficult to find a plausible identification strategy. This analysis provides important implications for broader economic theory with respect to intertemporal labor supply behavior.

The rest of the paper is structured as follows. Section 2 provides background on the physician work-hour regulations in the United States. Section 3 discusses the conceptual framework. Section 4 describes the data and shows the effectiveness of the regulations. Section 5 presents the empirical strategy, the identification, and the mean and distributional estimated results. Section 6 addresses potential mechanisms of the effects. Section 7 concludes.

2. Physician Work-Hour Regulations in the United States

Medical residency training traditionally requires lengthy work hours, but there was no regulation limiting the number of hours that could be assigned to a resident physician in the United States before the late 1980s. The public and the medical education establishment started to be aware of and investigate the consequences of overwork by residents after the death of an

18-year-old college freshman, Libby Zion, in 1984. As a result of the investigation, New York State adopted the recommendations by the committee that evaluated the training and supervision of physicians in the state, and enacted the Libby Zion Law in 1989. The law forbade residents in New York State hospitals to work more than 80 hours per week or 24 consecutive hours. This law became the first regulation in the nation that restricted hours worked by residents. However, most residency programs in New York were found in violation of the law ten years after its implementation (Wasserman, 2018), and thus the law was likely not adequately enforced.

In June 2002, the Accreditation Council for Graduate Medical Education (ACGME) granted preliminary approval to similar regulations for all residents working in accredited medical training institutions in the U.S.,¹ and the regulations were implemented in July 2003 (Philibert et al., 2002). With the aim to improve patient safety by reducing fatigue-related medical errors made by residents, the ACGME's standards consist of (1) a maximum of 80 hours worked per week, averaged over four weeks; (2) a 24-hour limit on continuous duty with an additional six hours allowed for patient transfer, administration, and didactic lectures; (3) one day in a week free of all medically related duties; (4) a limit on call frequency; (5) a 10-hour rest period between duty periods or work shifts; (6) a maximum of eight hours in the

¹ All the residency programs for doctors with a Doctor of Medicine (MD) degree and the majority of the programs for doctors with a Doctor of Osteopathic Medicine (DO) degree in the United States are ACGME-accredited. As of 2015, 7.6 percent of the practicing physicians hold a DO degree, and 65 percent of DO graduates enrolled in ACGME-accredited programs.

weekly hours allowed if it is beneficial for resident education, upon approval by the sponsoring institution and the Residency Review Committee of the specialty. In order to comply with the policy, many residency programs changed rotation schedules, decreased the call frequency, and replaced resident services with care by physician extenders (Philibert et al., 2009). With monitoring through program audits and periodic surveying of residents, penalties for non-compliance with the regulations included residency program probation and potential loss of accreditation.²

In 2008, the ACGME proposed minor revisions to the duty hour standards in response to the recommendations of the Institute of Medicine (IOM). The changes were made to a 16-hour limit on continuous duty for first-year residents. Residents in their second year and beyond followed the 24-hour limit with a reduction in additional hours for hand-offs from six hours to four hours. No changes were made to the 80-hour limit and call frequency. These new standards went into effect in July 2011. Though these standards were designed to improve patient safety by reducing residents' fatigue, they had also led to unintended negative consequences on residents' attainment of clinical skills. In March 2017, the ACGME further announced a policy change which raised the maximum number of consecutive hours from 16 to 24 hours for first-year residents, and this new standard went into effect in July 2017 (Asch et al., 2017).

² According to the investigation one year after the reform, five percent of the 2,235 programs that ACGME reviewed were found in violation of one of the standards. From the survey of 25,176 residents, 3.3 percent reported working more than 80 hours per week during the past four weeks.

There is an extensive literature on the effects of the ACGME regulations on residents' well-being and learning, as well as patients' safety and health outcomes. Most of the literature compares the outcomes before and after the reform in 2003 and uses observational cohort analysis from a single site, multiple sites, or national databases. Overall, the findings suggest that residents' well-being is improved between the pre- and post-2003 time periods. Many studies show that residents' fatigue has decreased since the implementation of the ACGME regulations (Gopal et al., 2005; Barrack et al., 2006; Hutter et al., 2006; Martini et al., 2006; Landrigan et al., 2008; Philibert et al., 2013), and some studies find that physicians are having more children, spending more time to attend family events, and leading less stressful lives since 2003 (Karamanoukian et al., 2006; Jones and Jones, 2007). However, the effects on residents' educational outcomes, patients' safety, and their health outcomes vary across studies, and these consequences may also differ between medical and surgical specialties.³

Despite numerous studies on the reform, little is known about its effects on physicians' employment patterns in the long run. To my knowledge, the only paper directly related to labor supply effects of the ACGME regulations is Wasserman (2018), which focuses on changes in residents' specialty choice. She finds that female physicians are more likely to enter a specialty when the specialty reduces its time requirements due to the reform, but there is little change in specialty entry response among male physicians. While the regulations reduce physicians' early

³ See Philibert et al. (2013) for a systematic literature review.

career hours, the question of interest is whether their long-term labor supply decisions are affected or not. The following analysis estimates the effects of the reform on physicians' post-residency hours of work, as a measure of long-term labor supply, and addresses potential mechanisms of the effects.

3. Conceptual Framework

To link this short-term policy to longer-term impact on labor supply, there are several distinct features of physician career paths that need to be taken into account. A physician needs to complete three to seven years of residency training after college and medical school, and then obtain medical licensure to practice medicine. Residency programs are typically run by hospitals and have a limited number of residency slots each year. Residents are paid on an annual basis, which is largely funded by the government, and there is little difference in resident salary across specialties. Since residents are forced to work a set of hours, their labor supply can be considered as perfectly inelastic within specialties. Therefore, physicians are likely to have less leisure time during residency because they are constrained to work more hours than they would otherwise choose. Despite the number of hours worked during residency, their average hours typically decrease after completing the training. The wage increases dramatically after residency as most practicing physicians make considerably more than resident salary, and post-residency salary varies largely across specialties.

As mentioned above, residents' labor supply choices are constrained by institutional rules regulating labor time and effort provision, and they are not free to choose their working hours. Therefore, regarding labor supply decisions in a life cycle framework, the intertemporal substitution hypothesis with time separable utility does not fit in this context. Alternatively, a theoretical hypothesis that can be used to explain how the reform may affect long-term labor supply is the neoclassical model with non-separable utility (Fehr and Goette, 2007). Holding the wage constant, this model predicts that an increase in a worker's effort in the previous period causes a higher disutility of labor in the following period, which decreases the worker's labor supply. Since the ACGME regulations cause an anticipated transitory reduction in physicians' labor supply in early career without changing their wage and lifetime wealth, the regulations decrease their disutility of effort during residency, and subsequently increase their labor supply in the long run, based on the prediction of this model.

An opposite theoretical hypothesis is the "persistence hypothesis" which states that individuals' work experience in early career is a major determinant of subsequent labor supply due to human capital accumulation, change in family commitments, and taste for work, among others (Clark and Summers, 1982). On human capital grounds, those who work more tend to accumulate more human capital, which in turn increases the return to work relative to leisure in the future (Heckman and Willis, 1979; Freeman, 1980; Clark and Summers, 1982). Those with a lack of work experience, on the other hand, may develop family commitments (such as

children) which reduce the return to work relative to staying outside of the labor force (Angrist and Evans, 1998; Sasser, 2005, Kleven et al., 2018). In addition, individuals' taste for work could be affected by prior work experience according to habit formation effects (Clark and Summers, 1982; Clark, 1999).

This hypothesis suggests that a short-term reduction in physicians' early career hours tends to persist after residency. Intuitively, human capital accumulated through residency experience affects labor supply in the future. The literature finds that the work-hour limits reduce continuity of care and educational continuity for residents (Feanny et al., 2005; Vidyarthi et al., 2006; McBurney et al., 2008; Nakayama et al., 2009). These losses of perioperative continuity may lead to negative consequences for residents' professional development and preparedness for practice, especially in surgical specialties (Philibert et al., 2013). Since there are substantially more male physicians in surgical specialties, they experience the greatest reduction in resident hours as well as human capital accumulation. The negative impact of the reform on long-term labor supply for male physicians may be largely explained by this potential mechanism.

Other potential mechanisms pertain to marriage and fertility decisions. Since the work-hour regulations affect residency training, which occurs during the prime childbearing years, physicians would plan the timing of marriage and fertility relative to their residency. Theoretically, the regulations alter the labor market conditions for residents and thus change

their opportunity cost of time factored into the fertility transition. Previous research has shown that the reform results in physicians having more children, spending more time to attend family events, and leading less stressful lives (Karamanoukian et al., 2006; Jones and Jones, 2007). Consequently, these changes in family commitments, especially childbearing, may decrease their labor-force attachment and keep them from developing further their careers.

The dynamic impact of children on labor market outcomes also greatly depends on spousal income (Goldin, 2014). In two-income households, if their partner is doing well financially, they may feel more comfortable pulling back on their hours. As such, physicians with higher-earnings spouses have a lower opportunity cost of career interruptions (Sarma et al., 2011). According to the AMA Masterfile, nearly 40 percent of physicians marry another physician or health care professional. In addition, most of the female physicians are married to male physicians, while the reverse is not true (Sasser, 2005). With higher-earnings physician spouses, who also work long hours, new physician mothers face more binding constraints on hours and a lower opportunity cost of career interruptions. Therefore, they are more likely to reduce their post-residency hours than male physicians, who are less likely to have physician spouses. In Section 6, I provide empirical evidence on the mechanisms pertaining to marriage and fertility across gender.

Overall, the neoclassical model with non-separable utility predicts that the ACGME regulations decrease physicians' disutility of work during residency and thus increase their

labor supply in the long run. In contrast, the persistence hypothesis suggests that the regulations reduce resident hours and at the same time lower the opportunity cost of work time, especially for those who would have worked more than the work-hour limits in the absence of the reform. This could lead to less human capital, more family commitments, and habit formation effects for those exposed to the reform, and thus results in a reduction in hours worked over time. From the above discussion, the potential impact of the ACGME regulations on long-term labor supply is ambiguous due to the contradicting effects between these two hypotheses; therefore, empirical evidence is needed to better understand physicians' employment patterns. As it will be presented in Section 5, my empirical results are consistent with the persistence hypothesis and suggest that the short-run reduction in labor supplied persists even when physicians are not bound by the work-hour limits.

4. Data and the Effectiveness of the Reform

4.1 Data Construction and Summary Statistics

This analysis uses data from the monthly Current Population Survey (CPS) between 1989 and 2017. Administered by the U.S. Census Bureau, the monthly CPS is a household-based survey which selects a nationally representative sample and contains a large amount of demographic and employment information. To identify physicians in the CPS, I restrict the sample to the civilian non-institutional population who hold an advanced degree and reported

their occupation as a “physician or surgeon.” To ensure sufficient attachment to the labor market while still allowing part-time work, I limit the analysis to active physicians who are currently in the labor force and work more than 20 hours per week.⁴

Whether a physician was exposed to the work-hour regulations is based on the year of residency training, but such information is not available in the CPS. Inspired by Staiger et al. (2010), I identify physicians as residents if they were younger than 35 and use the year of birth as a proxy for exposure to the ACGME regulations.⁵ Physicians who could have been potentially subjected to the regulations were born after 1968 (i.e., those who worked as a resident after 2003), and they are categorized as the treatment group. For physicians trained in New York, although they might have been potentially exposed to similar regulations, the Libby Zion Law, most residency programs in New York were found in violation of the law (Wasserman, 2018). In addition, I use the CPS Annual Social and Economic Supplement (ASEC) data to test whether there was a significant decrease in resident hours around the implementation of the law in 1989.⁶ Figure 1 shows that the average hours worked by residents remained fairly stable around 1989 and changed little through the 1980s and 1990s, suggesting that the law was inadequately enforced.⁷ On the other hand, the average hours worked by

⁴ The American Medical Association (AMA) defines “active” as working more than 20 hours per week.

⁵ According to the 2007 AMA Physician Masterfile data, which is the primary source of physician workforce data in the U.S., Staiger et al. (2010) point out that 97% of hospital-based physicians younger than 35 were residents. However, not all residents were trained in a hospital-based program, and thus using age 35 to identify residency status might lead to a potential source of bias. This problem is addressed in Section 5.3.

⁶ The monthly CPS does not provide the hours worked variable before 1989, which is the main reason why the analysis period starts from 1989.

⁷ I also look at the trend for New York only and find no significant change in hours worked by residents around 1989 either. However, the sample size of New York physicians in the ASEC is very small (around 28 observations

residents decreased sharply following the imposition of the work-hour limits in 2003, demonstrating that the ACGME regulations effectively led to hours cut. Therefore, the central variation in the empirical analysis below comes from the cohort variation in exposure to the 2003 ACGME reform.

According to the U.S. Department of Health and Human Services, more than 95 percent of the physicians graduated from medical school after age 26. In addition, by the end of the sample period in 2017, the oldest possible age that the treatment group can achieve is 48. Therefore, the analysis focuses on active physicians aged 26 to 48 with non-missing values for weekly hours worked.⁸ The analysis sample comprises 69,371 physicians. Throughout this paper, data are weighted using CPS sampling weights and adjusted for respondent-level clustering to take into account the rotating panel design of the monthly CPS. It is worth noting that this selected group are at their prime working and childbearing ages, which helps understand the role of job flexibility in the work-family interface.

Table 1 shows the summary statistics of this analysis sample. The mean age of the sample is 37.6, and approximately 66.5 percent of the subjects are male. The average hours worked per week is 55.5 (standard deviation = 17.4). With respect to gender differences, male physicians tend to be slightly older and have smaller proportion of the treated population, due

per year on average), and thus it is not informative enough.

⁸ The measure of hours worked is based on the self-reported hours in the previous week in the monthly CPS. An alternative measure is the usual number of hours per week (over an unspecified time period), but it is not available until 1994 in the monthly data. I chose the former for the analysis since it has a relatively shorter-term recall and is available for a longer period of time. Note that the hours worked measure is top-coded at 99 hours prior to 1994; however, there are only 690 observations (about 1%) in the sample at the top-coded value of 99 before 1994.

to the increasing female share of physicians in recent decades. In addition, male physicians work about six hours more than their female counterparts, and they have a higher rate of marriage and have more children on average. With respect to differences by treatment status, the treatment group is older and comprises more females. The average hours worked per week is 55.3 for the treatment group and 55.8 for the control group.

There are many advantages of using the monthly CPS data for this analysis. First, it provides repeated cross-sectional observations over a longer time period than any other comparably sized dataset that includes physicians' information. This is particularly useful for analyzing the effects on lifecycle patterns. The large enough sample also helps conduct analyses separately by subpopulations and run robustness tests using different regression specifications. Second, the CPS includes important demographic characteristics and employment information. These variables matter for identification because they allow us to control for dynamic changes in the composition of the physician workforce that might cause potential imbalances in the demographic characteristics between the treatment and control groups. Third, the CPS data are more up-to-date than the American Medical Association (AMA) Physician Masterfile in terms of physicians' employment information and can be used as a benchmark dataset on physician labor supply (Staiger et al., 2009). Fourth, it is widely conjectured that residents' self-reported hours from the ACGME monitoring data may underestimate hours worked due to the desire to protect residency programs or pressure from

residency program directors (Landrigan et al. 2006; Szymczak et al. 2010; Fargen and Rosen, 2013). The CPS data, collected by non-ACGME researchers, limit the potential for this type of misreporting.

Nevertheless, the CPS lacks physician-specific information regarding residency training and specialty choice. Therefore, I cannot directly identify the treatment status and analyze some of the other interesting labor market outcomes. In addition, the information on income is top-coded in the CPS for confidentiality reasons. More than 85 percent of the sample analyzed here has top-coded values or missing values in weekly earnings at ages 35 to 48.⁹ For these reasons, it is difficult to provide direct evidence on how the regulations change physicians' earnings profiles over time.

4.2 First Stage: The Effectiveness of the Reform

The trend in resident hours in Figure 1 provides preliminary evidence that the 2003 reform was effective in reducing resident hours on average. To assess the effectiveness of the regulations for the analysis, I plot the trends in hours worked by resident and non-resident physicians using the analysis sample in Figure 2. Prior to 2003, residents were not exposed to the work-hour limits. As shown, the average hours worked per week by residents remained high through 2002 and then declined sharply after the preliminary approval of the ACGME

⁹ Total weekly earnings are top-coded at \$1923 prior to 1997 and \$2885 since 1998. For hourly wages, more than 90 percent of the physicians in the CPS have missing values.

reform in 2002 and its implementation in 2003. The average hours worked per week decreased from 64 in 2002 to 58 in 2004. This sharp decline in average resident hours after the introduction of the reform provides evidence that the regulations were indeed enforced.

In contrast, such plummet was not found in the work hours trend of nonresident physicians since they were not restricted by the regulations. Instead, their hours worked have gradually trended downward since 1999. This trend can be largely explained by the aging of the physician population and the increasing proportion of female physicians, who practice fewer hours than their male counterparts on average (Crossley et al., 2009; Staiger et al., 2010; Sarma et al., 2011, Wang and Sweetman, 2013). Staiger et al. (2010) also point out other possible factors that drive the downward trend of hours worked by practicing physicians, such as the decrease in physician fees since the early 1990s, developments among both public and private payers in the 1990s, and improvements in physician productivity due to technology. In addition, the rapid growth in care by hospitalists in the late 1990s and the early 2000s (Kuo et al., 2009) and the HMO penetrations in the 1990s (Zhan et al., 2004) may also be attributed to the decreasing hours worked by practicing physicians.

Figure 3 compares the trends in hours worked between physicians who were exposed to the regulations and those who were not. As a result of the reform, the treatment group (physicians born after 1968) worked fewer hours per week than the control group (physicians born before 1968) during residency (ages 26 to 34). Interestingly, the difference in hours

worked between the treatment and control groups still remained even when physicians were not constrained by the work-hour limits after residency (above age 35), showing a persistent decline in hours worked.

In addition to visual evidence, I estimate the effect of the reform on resident hours for the full sample and by gender. Since residency programs must comply with all the work-hour standards, they should be considered as a whole when interpreting the estimated effects. The empirical specification is as follows:

$$Y_i = \beta_0 + \beta_1 \text{Exposed}_i + X_i \beta_2 + \alpha_c + \alpha_t + \alpha_s + \varepsilon_i, \quad (1)$$

where Y_i is the weekly hours worked by physician i aged 26 to 34. Exposed_i is a dummy variable indicating exposure to the ACGME regulations during residency training, which equals 1 for physicians born after 1968 and 0 for physicians born before 1968. X_i is a set of covariates including age, gender, and race/ethnicity. α_c , α_t , and α_s denote cohort, year, and state fixed effects, respectively. The estimates of β_1 identify the effects of the reform on resident hours and are shown in the first row of Table 2. As a result of the reform, the mean resident hours per week significantly decreased by 8.4 overall, which supports the effectiveness of the ACGME reform. Analyzing by gender, the reform reduced the mean resident hours per week by 10.54 for males and 5.94 for females, but the estimate for females is imprecisely estimated.

There are two main possible reasons for the insignificance of the effect on resident hours for female physicians. First, there are fewer female physicians whose hours were capped by the regulations. In my analysis sample, only the top 20 percent of the female control group exceeds 80 hours worked per week. Wasserman (2018) also shows that female physicians were less likely to choose the most time-intensive specialties where the hours worked by residents were more than 80 hours per week before the reform. Hence, the majority of the female physicians were not primarily affected by the reform, and the mean estimate of the effect on resident hours is attenuated. Second, females are found to enter more time-intensive specialties as a result of the reform, whereas there is little change in males' specialty choice (Wasserman, 2018). If females change their specialty choice in response to the reform, they are also potentially altering their residency hours towards longer hours. This behavioral change increases mean hours worked by female residents and balances out their effect of the reform on resident hours which were originally designed to reduce hours worked.

5. The Impact on Long-Term Labor Supply

5.1 Difference-in-Differences Approach

Overall, Figures 2 and the estimated effects of the reform on resident hours provide empirical evidence on the effectiveness of the ACGME regulations, and Figure 3 suggests that the treatment group has fewer hours worked after residency relative to the control group. A

fundamental challenge in interpreting this pattern as causal is that the cohort variation that identifies differences in exposure to the regulations is time-series in nature. Omitted variables that are correlated with changes in labor supply and the exposure, and other secular trends that affect physician hours, might explain this pattern as well, resulting in an identification problem.

To tackle this problem and identify the causal impact, I begin by the baseline difference-in-differences approach using the cohort-time variation in exposure to the ACGME regulations. Since the reform affects residents trained after 2003, the strategy is to compare the change in hours worked from residency to post-residency between the treatment and control groups. The regression framework of the baseline model is as follows:

$$Y_{ict} = \beta_0 + \delta_1 \text{Exposed}_c + \delta_2 \text{Post}_t + \delta_3 (\text{Exposed}_c \cdot \text{Post}_t) + \varepsilon_{ict}, \quad (2)$$

where i indexes physicians, c indexes birth cohorts, and t indexes years. The outcome variable Y_{ict} is the weekly hours worked by physicians. Similar to the definition in Equation (1), Exposed_c is an indicator of exposure to the ACGME regulations during residency training. Post_t is an indicator of completing residency in year t , which equals 1 for physicians aged 35-48 and 0 for physicians aged 26-34. δ_3 is the coefficient of interest. To identify it as the casual impact, the treatment and control groups are assumed to have the same trends in hours worked over time in the absence of the regulations.

Since the composition of the physician workforce has changed in recent decades, especially with the female share of physicians increasing substantially, there are some factors

that affect labor supply and may also be correlated with exposure to the regulations. I control for pre-treatment baseline observables to account for potential imbalances in the demographic characteristics between the treatment and control groups. I also expand the model by including cohort and year fixed effects to control for additional unobserved factors. The empirical specification can be written as:

$$Y_{ict} = \beta_0 + \alpha_c + \alpha_t + \beta_1(\text{Exposed}_c \cdot \text{Post}_t) + X_{ict}\beta_2 + \alpha_s + \varepsilon_{ict}. \quad (3)$$

X_{ict} is a set of pre-determined demographic controls including age, gender, and race/ethnicity. α_c reflects fixed effects for birth cohorts, and α_t reflects fixed effects for the calendar year in which they are observed. The cohort fixed effects control for differences across cohorts in the outcome variable, and the year fixed effects control for any general time trends in the outcome variable, picking up some of those possible factors that drive the long-term secular decline in physician work hours mentioned in Section 4.2. I also add a set of state dummies α_s to control for any time-invariant unobservables that affect the outcome variable across states. In particular, it accounts for the state differences in the institutional design features, such as state-specific licensing requirements. I do not include controls which may cause potential endogeneity with respect to labor supply (e.g., family formation and practice setting). The coefficient β_1 identifies the treatment effect, by contrasting the hours worked from residency to post-residency between physicians who were exposed to the regulations and those who were not.

Previous studies have documented differences in labor market outcomes between male and female physicians (e.g., Rizzo and Blumenthal, 1994; Sasser, 2005; Rizzo and Zeckhauser, 2007; Wang and Sweetman, 2013; Wasserman, 2018). In addition, my regression results (shown in Section 5.2) also indicate that gender has a significant impact on long-term labor supply. In addition to the analysis on the full sample, I also estimate the effects by gender to explore whether the ACGME regulations affect male and female physicians differently.

The key identifying assumption in Equation (3) is the so-called parallel trends assumption, that is, the evolution of hours worked between the treatment and control groups (conditional on observed variables) is the same over time in the absence of the reform. This analysis has several advantages in meeting this condition. First, the ACGME regulations were enacted in order to reduce fatigue-related medical errors made by residents. The motivation and the nature of this policy make it unlikely to be correlated with other policies that affect physician labor supply or costs of childrearing. Since the model includes cohort and year fixed effects, and a set of demographic controls, the estimated effects on the outcome variables can be attributed to changes in hours worked during residency training within cohorts over time. Second, the sample is fairly homogeneous with respect to skills and other traits. Since this analysis is based on a profession that has highly competitive entry requirements, rigorous educational standards, and very specialized training, the general concern about unobserved

heterogeneity across individuals or cohorts is considerably diminished. Altogether, these ameliorate the threat from unobserved factors to confound the estimated effects.

5.2 Estimated Mean Effects

Table 3 shows the estimation results using the baseline difference-in-differences model, specified in Equation (2), for the overall sample and by gender. Each cell contains the mean hours worked for its subgroup of the sample. For the overall sample shown in Panel A, the residency versus post-residency difference in hours is 8.91 for the treatment group and 6.93 for the control group. Thus, the treatment group worked 1.98 less hours per week than the control group from residency to post-residency. With a standard error of 0.66, it is statistically different from zero. The analysis by gender is shown in Panels B and C. Among male physicians, the treatment group decreases weekly hours worked by 1.11 (or 2.25 percent) more than the control group, but the estimate is statistically insignificant. Among female physicians, the treatment group decreases weekly hours worked by 1.86 (or 3.29 percent) more than the control group. The results suggest that the long-term labor supply effect for female physicians is greater than that for male physicians, but they are not statistically different (p-value = 0.59).

Controlling for observed demographic differences between groups and the cohort and year fixed effects, the preferred estimates of the long-term labor supply effects using Equation (3) are shown in Table 4. Columns (1) and (2) display the results for the full sample, without

and with the inclusion of demographic controls, respectively. The first row reports the estimated coefficients of interest. The estimate in Column (2) implies that the reform decreases physicians' mean hours worked at ages 35 to 48 by roughly 4.23 hours per week, which is a statistically significant 7.75 percent decrease over the control group mean of 54.55 hours. Columns (3) to (6) show the estimates by gender. The results indicate that the regulations reduce both male and female physicians' long-term hours worked by roughly four hours per week on average, and the effects are not statistically different across gender (p-value = 0.67). These findings are robust to the inclusion of the covariates. The next six rows in Columns (2), (4), and (6) report the estimated coefficients of the demographic characteristics on long-term hours worked. Among all three columns, the average hours worked decrease significantly with age, and as shown in Column (2), gender seems to have a large and significant impact on post-residency hours, with males working more than females by 6.34 hours per week.

Since men tend to enter specialties that require longer hours worked, there are more male physicians whose hours were capped by the regulations. To obtain the effects of a given reduction in hours during residency on subsequent labor supply, I take into account the effects of the reform on resident hours and divide the reduced form coefficients in Columns (2), (4), and (6) by their corresponding first-stage estimates. For the full sample, a reduction in hours during residency decreases labor supply at ages 35 to 48 by 0.5 hours per week on average (standard error = 0.02). Analyzing by gender, a reduction in hours during residency results in a

decrease of 0.37 hours (standard error = 0.02) for males and 0.75 hours (standard error = 0.07) for females after residency.¹⁰ Although female physicians in general were less likely to be restricted by the work-hour limits, these findings suggest that females are more responsive to a given reduction in early career hours caused by the reform, and this gender difference is statistically significant at the 0.01 level.

There are a few observations with likely unreliable self-reported hours in the data.¹¹ To address this concern and reduce the impact of possibly spurious outliers, I repeat the analysis using winsorized and trimmed hours worked at the 1% and 99% levels, as well as the 5% and 95% levels, as alternative outcome variables. Winsorizing at the 1% and 99% levels sets the bottom 1% to the 1th percentile and the top 1% to the 99th percentile. Without discarding the extreme values, the winsorization method still takes those values into account and treats them as if respondents reveal certain information on their hours worked. The estimated effects using these alternatives are very similar to the results shown above.¹²

5.3 Potential Threats to Identification

Although the inclusion of control variables and the homogeneity of the physician workforce ameliorate potential threats from unobserved confounders, there may still be at least

¹⁰ These estimates are obtained by $-4.226/-8.399 = 0.503$ for the full sample, $-3.883/-10.538 = 0.369$ for males, and $-4.431/-5.939 = 0.746$ for females.

¹¹ The maximum of self-reported hours worked is 192 in the data, which is obviously exaggerated.

¹² The results are available upon request.

two other concerns regarding the identification. The first problem pertains to the year-of-birth proxy for treatment status. Ideally, I would use the year of residency to identify individual exposure to the ACGME reform, but the CPS does not contain such information. This may result in misclassifications of the treatment and control groups. The second problem pertains to any remaining unobserved heterogeneity of exposure to the regulations with respect to labor supply. I discuss below these identification issues and how I attempt to assess them.

First, using age 35 to identify residency status is a potential source of bias. Depending on the medical specialty, the length of residency training ranges from three to seven years.¹³ This leads to some variation in the age that a physician can complete residency.¹⁴ There are two possible misclassifications of the treatment status: (1) physicians who were born after 1968 and completed residency training before 2003, and (2) physicians who were born before 1968 and completed residency training after 2003. In the first case, the treatment group contains non-treated individuals; in the second case, the control group contains treated individuals. As a consequence, the magnitude of the estimated effects would be underestimated in both cases.

To assess the effects of potential misclassifications, I estimate three alternative specifications which reclassify the treatment and control groups with tighter year-of-birth windows, and the age proxies for residency status are also adjusted accordingly. Since the main

¹³ For example, internal medicine, general surgery, and neurosurgery require three, five, and seven years of training, respectively.

¹⁴ Nearly all physicians graduate from medical school after age 26. With a minimum of three years for residency training, the earliest possible age to complete residency is 29. Similarly, with a maximum of seven years for residency training and allowing a gap of five years at some point, the oldest possible age to complete residency is likely to be 38.

analysis uses physicians born before and after 1968 as the control and treatment groups, respectively, the alternative specifications adjust the year-of-birth proxies for the treatment status as follows: (1) the treatment (control) group consists of physicians born after 1969 (before 1967); (2) the treatment (control) group consists of physicians born after 1970 (before 1966); (3) the treatment (control) group consists of physicians born after 1971 (before 1965). When using tighter year-of-birth windows for the treatment status, there should be less misclassification; however, it discards non-negligible amount of observations. As shown in Table A1 in the Appendix, the results using these alternative treatment and control groups are similar to the main results.

Second, there may still exist some unobserved factors that cause nonrandom selection of individuals into the physician workforce following the implementation of the reform. It could be that the reform changed the pool of applicants and entrants into the medical profession in dimensions not captured by the admission criteria and observed characteristics, but are relevant to the labor market. Specifically, given the decline in hours requirements during residency, individuals who prefer balanced lifestyles would be induced to enroll in medical school. These unobserved preferences are correlated with a priori disposition toward fewer hours worked at later ages, which would lead to a decrease in the average hours worked by physicians over time. As a result, the magnitude of the negative effects of the reform on long-term labor supply would be overstated.

To assess whether the effects are driven by this selection bias, I estimate an alternative specification, taking as the treatment group physicians who already entered medical school at the time of the reform but were trained under the new regulations during residency. Given that the ACGME regulations were approved in 2002 and that the fresh college graduates in 2002 were born in 1979-1980, I restrict my sample to the 1941-1980 cohorts and re-estimate the effects. Since there may be some physicians enrolling in medical school few years after college, I also restrict the sample to the 1941-1978 cohorts and the 1941-1976 cohorts to test the robustness of my results. As shown in Table A2 in the Appendix, the key estimates are robust to these alternative specifications, suggesting that the likelihood of this self-selection confounding the results is not considerable.

Although unobserved heterogeneity seems less applicable to physicians who have been through highly competitive admission process and invested many years in formal training beyond college with the completion of medical school and residency training, the above robustness check suggests that remaining unobservable heterogeneity is no a significant concern. Other empirical studies also find no evidence of reduced hours worked driven by unobserved preferences for balanced lifestyles among younger physician cohorts (Crossley et al., 2009; Staiger et al., 2010; Sarma et al., 2011).

5.4 Distributional Effects

According to the ACGME work-hour standards, the regulations should primarily affect residents who would have worked more than the work-hour limits (e.g., 80 hours per week) in the absence of the reform. This naturally leads to a disproportionate impact on those at the upper tail of the hours distribution. As shown in Table A3 in the Appendix, only those above the 75th percentile among the control group aged 26-34 (unaffected residents) work over 80 hours per week. The estimated mean impacts may be attenuated by the other part of the distribution and incompletely reveal the effects on those affected.

To identify the heterogeneous impacts across the hours distribution, I use a changes-in-changes (CIC) model following Athey and Imbens (2006) and Melly and Santangelo (2015) that extends the model to include covariates. The CIC framework relaxes the parallel trends assumption and provides unconditional treatment effects of the whole distribution, which fits with the study of interest. The estimated quantile treatment effects provide evidence on what would happen to the overall hours distribution in the long run if there is a policy regulating the maximum workweek during residency. I estimate CIC effects for 17 quantile values, $q = \{0.1, 0.15, \dots, 0.85, 0.9\}$, and their bootstrapped 95-percent confidence intervals with 1,000 replications, controlling for age, gender, race/ethnicity, and cohort and year fixed effects.

Figure 4 shows the CIC estimates for the full sample (Panel A) and by gender (Panels B and C). In general, the effects become negative and stronger when moving towards the upper

end of the work hours distribution. As shown in all three panels, the quantile treatment effects are not statistically different from zero between the 10th and the 70th quantiles, and the effects become negative and larger than the mean estimates at the top of the distribution. This finding confirms that the reform primarily affects those with the highest hours of work. For the full sample, the estimates above the 80th quantile are statistically different from zero and remain stable at around -5.56. For male physicians, the estimates above the 75th quantile are statistically significant and have an average level of -8.58. For female physicians, the estimates are imprecisely estimated, but overall negative and hovering around -4.5 for those above the 70th quantile (except for the 80th quantile). At the extreme upper tail of the distribution (the 90th quantile), the magnitudes of the effects for male and female physicians are similar.

There are two potential reasons why the effects are greater for males than for females at most of the upper quantiles. First, since hours worked by male physicians at the upper quantiles are more likely to be capped by the regulations, the reform reduces their long-term labor supply more than their female counterparts. Second, more female physicians enter the most time-intensive specialties as a result of the reform, whereas there is little change in specialty choice by male physicians (Wasserman, 2018). If more female physicians enter long-hours specialties in response to the reform, they are also potentially altering their residency hours and increasing hours worked after residency. The link between changes in specialty

choice and hours worked provides a potential explanation and mechanism for the gender difference in the effects of the reform on the distribution of hours worked.¹⁵

6. Mechanisms

The work-hour regulations reduce physicians' early career hours and at the same time lower the opportunity cost of work time. As discussed in Section 3, this could lead to less human capital, more family commitments, and change in taste for work for the treatment group. Therefore, the reform results in a negative impact on physicians' long-term labor supply. Since the regulations affect residency training, which occurs during the prime childbearing years, the mechanisms pertaining to marriage and fertility decisions are particularly of interest. Various studies suggest that home production is disproportionately undertaken by females even within this highly skilled profession (Gronau, 1988; Coleman and Pencavel, 1993; Angrist and Evans, 1998; Sasser, 2005; Goldin, 2014; Kleven et al., 2018). Women may choose the specialty and work environment that are family friendly, and avoid jobs with long hours and greater career advancement possibilities. Wang and Sweetman (2013) show that married female physicians work fewer hours per week than both their married male counterparts and their unmarried

¹⁵ There are two caveats of these estimated quantile treatment effects. First, the existence of point masses in the hours worked data might contaminate the effects, and thus the results must be evaluated with caution. Second, whether men or women are more responsive to a given reduction in resident hours across the hours distribution is unknown without knowing the corresponding distributional effects of the reform on resident hours.

female counterparts. The impact of children on women’s labor market outcomes is large and persistent, whereas there is little evidence on such impact on men.

The CPS data allow us to learn about the presence of marriage and fertility mechanisms in the context of the ACGME reform and whether there are gender differences in these mechanisms. I use the following regression framework to investigate the impact of exposure to the regulations on male and female physicians’ marriage and fertility outcomes:

$$Y_i = \beta_0 + \beta_1 \text{Exposed}_i + X_i \beta_2 + \alpha_t + \alpha_s + \varepsilon_i. \tag{4}$$

When examining the marriage mechanism, I use binary indicators of marriage and divorce for physicians between ages 26 and 48 as the outcome variables. For the effects on fertility decisions, I examine two outcomes: (1) fertility at the extensive margin, which is an indicator of having at least one child; (2) fertility at the intensive margin, which is the total number of children (completed fertility).¹⁷ By contrasting these outcomes between physicians who were exposed to the regulations and those who were not, the estimates of β_1 show the effects of interest. The identification is based on the selection-on-observables assumption; that is, there is no unobserved factor that affects both outcomes (marriage and fertility) and treatment (exposure to the reform). This assumption plausibly holds since the reform is not correlated with other policies that would affect physicians’ marriage and childrearing. The set of control

¹⁶ Note that this is not a difference-in-differences model in which its estimates may only reflect how the reform changes the timing of marriage and fertility. Instead, I use the regression framework specified in Equation (4) to capture the total impact of the reform on marriage and fertility outcomes.

¹⁷ According to the National Association for Public Health Statistics and Information Systems, completed fertility is defined as the number of children to a person by the end of a woman’s childbearing years, 15 to 44 years old, the latest age at which people typically have their last child.

variables and the homogeneity of the physician workforce also make the identification assumption more plausible to be satisfied.

Table 5 shows the effects of the work-hour regulations on the likelihoods of marriage and divorce between ages 26 and 48, controlling for demographic characteristics (age, gender, and race/ethnicity), year fixed effects, and state fixed effects. Columns (1) and (2) show the results using the full sample, and Columns (3) to (6) present the effects by gender. The estimated coefficients of interest, shown in the first row, suggest that the reform significantly increases the probability of marriage for female physicians by 9.2 percentage points, but there is no significant impact on male physicians' marriage rate. There is no change in divorce rates for both male and female physicians. These estimates suggest that marriage decisions made by females are more sensitive to the work-hour regulations than those made by males.

Table 6 presents the estimated effects of the regulations on physicians' fertility between ages 26 and 48. Columns (1), (3), and (5) show the effects on fertility at the extensive margin for the full sample and by gender. The estimates suggest that the reform increases the probability of having a child for female physicians by 9 percentage points, which is statistically significant at the 1 percent level. However, the reform does not affect the probability of having a child for male physicians. Columns (2), (4), and (6) show the estimated effects on completed fertility. The results also suggest a substantial gender difference. Exposure to the regulations leads to a significant increase of 0.2 children for female physicians, but it does not affect the

total fertility of male physicians. The estimation results using probit and logit models for the binary outcomes of marriage and fertility are very similar to the above results, as shown in Table A4 in the Appendix. In addition, I estimate the effects for physicians aged 26-34 and 35-48 separately, as shown in Table A5 in the Appendix. These estimates provide suggestive evidence on how the reform changes the timing of marriage and fertility beyond the total impact of the reform. The results indicate that the reform increases the marriage and fertility outcomes in the post-residency period more than those during residency, but they are not statistically different from each other.

To sum up, I find substantial gender differences in both marriage and fertility decisions in response to the reform. The results show that the regulations raise the likelihood of marriage and have positive and significant effects on fertility at both the extensive and intensive margins for female physicians. On the contrary, the regulations have little impact on marriage and fertility outcomes for male physicians. During the childbearing years, these effects for females may result in a potentially supply shift that accounts for at least some of the decrease in their long-term labor supply. As discussed in Section 3, most female physicians are married to physician spouses, who are likely to have high earnings and work long hours. As the marriage and fertility rates increase for female physicians due to the reform, they likely face greater household obligations and more binding constraints on hours with a lower opportunity cost of career interruptions. As a result, this situation would lead them to work fewer hours, more

regular schedules, and generally more conducive to combining career and family. The estimated results in this section provide empirical evidence to support potential mechanisms of females' labor supply responses to a policy that reduces time requirements during the prime childbearing years. Changes in family commitments likely account for females' greater responsiveness to a given reduction in hours. These findings are also consistent with previous studies on gender differences in the relationship between childbearing and labor market outcomes (Gronau, 1988; Coleman and Pencavel, 1993; Angrist and Evans, 1998; Sasser, 2005; Wang and Sweetman, 2013; Goldin, 2014; Kleven et al., 2018).

7. Conclusions

This paper estimates the impact of the work-hour regulations that limit the maximum hours worked by residents on physicians' long-term labor supply. As a result of the 2003 ACGME reform, the mean resident hours per week decrease by 10.54 for males and 5.94 for females, but the estimate for females is not statistically significant. Exploiting the cohort-time variation in exposure to the reform, I contrast the hours worked from residency to post-residency between the treatment and control groups using a difference-in-differences approach. The results suggest that the reform significantly reduces mean hours worked after residency by about four hours per week for both male and female physicians. When taking the effects of the reform on resident hours by gender into consideration, women seem to be more responsive to

a given reduction in early career hours caused by the reform.

The heterogeneous impacts across the hours distribution are revealed by a changes-in-changes model. The estimated quantile treatment effects show that the reform does not have a statistically significant impact on those below the 70th quantile of the hours distribution. When moving towards the upper tail of the distribution, the effects become significantly negative and larger than the mean estimates, which confirms that those at the upper end of the distribution are primarily affected by the reform. Since hours worked by male physicians are more likely to be capped by the regulations, the reform reduces their long-term labor supply more than their female counterparts at the upper quantiles. However, at the extreme upper tail of the distribution, the magnitudes of the effects for male and female physicians are similar.

To reveal potential mechanisms of the effects uncovered on long-term labor supply, I examine how the regulations affect physicians' marriage and fertility outcomes across gender. The empirical evidence suggests substantial gender differences in marriage and fertility choices in response to the reform that reduces work time requirements during the prime childbearing years. The work-hour limits significantly raise the probabilities of marriage and having a child, as well as the total number of children, for female physicians. By contrast, the regulations have little impact on marriage and fertility outcomes for male physicians. These results indicate that changes in family commitments could be potential mechanisms for females' long-term labor supply effects. On the other hand, since there are substantially more male physicians in surgical

specialties that suffer the greatest reduction in hours due to the reform, they may experience detrimental effects on their professional development and preparedness for practice. The mechanism of human capital accumulation might potentially account for males' long-term labor supply effects, although I am not able to explore this mechanism empirically.

With increasing participation of females in the physician workforce and limited evidence on long-term labor supply responses to a reduction in early career hours, the physicians' hours of work decisions and potential gender differences are substantive issues for human resource policies in the health care sector. With respect to the effects of the reform on long-term labor supply, the reduction of four hours per week among practicing physicians in younger cohorts does not seem small. Policy makers may need to take into account changes in the amount of patient care hours or services provided by practicing physicians when addressing the projected supply of healthcare. With respect to gender differences in marriage and fertility decisions, the results indicate that less time requirements may help women plan the timing of marriage and fertility relative to their residency. Since residency training occurs during the prime childbearing years, and nowadays almost half of the medical students are women, developing workplace policies to accommodate pregnancy and childbearing is important for the medical profession. Future research may consider using other physician surveys that include information on earnings and other dimensions of physicians' labor market outcomes. In particular, dynamic changes in earnings profiles caused by the regulations can give important

insight into labor supply behavior. The findings can also provide direct evidence of the resulting effects on the gender wage gap.

References

- Angrist, J. D. and Evans, W. N. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review* 88 (3): 450-477.
- Asch, D. A., Bilimoria, K. Y., and Desai, S. V. 2017. "Resident Duty Hours and Medical Education Policy – Raising the Evidence Bar." *New England Journal of Medicine* 376 (18): 1704-1706.
- Athey, S. and Imbens, G. W. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* 74 (2), 431-497.
- Barrack, R. L., Miller, L. S., Sotile, W. M., Sotile, M. O., and Rubash, H. E. 2006. "Effect of Duty Hour Standards on Burnout among Orthopaedic Surgery Residents." *Clinical Orthopaedics and Related Research* 449: 134-137.
- Buske, L. 2004. "Younger Physicians Providing Less Direct Patient Care." *Canadian Medical Association Journal* 170 (8): 1217.
- Chen, M. K. and Chevalier, J. A. 2012. "Are Women Overinvesting in Education? Evidence from the Medical Profession." *Journal of Human Capital* 6 (2): 124-149.
- Clark, K. B. and Summers, L. H. 1982. "Labour Force Participation: Timing and Persistence." *Review of Economic Studies* 49 (5): 825-844.
- Clark, A. E. 1999. "Are Wages Habit-Forming? Evidence from Micro Data." *Journal of Economic Behavior & Organization* 39 (2): 179-200.
- Coleman, M. T. and Pencavel, J. 1993. "Trends in Market Work Behavior of Women Since 1940." *Industrial and Labor Relations Review* 46 (4): 653-676.
- Crossley, T. F., Hurley, J., and Jeon, S. H. 2009. "Physician Labour Supply in Canada: A Cohort Analysis." *Health Economics* 18 (4): 437-456.
- Fargen, K. M. and Rosen, C. L. 2013. "Are Duty Hour Regulations Promoting a Culture of Dishonesty Among Resident Physicians?" *Journal of Graduate Medical Education* 5 (4): 553-555.
- Feanny, M. A., Scott, B. G., Mattox, K. L., and Hirshberg, A. 2005. "Impact of the 80-Hour Work Week on Resident Emergency Operative Experience." *American Journal of Surgery* 190 (6): 947-949.
- Fehr, E. and Goette, L. 2007. "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment." *American Economic Review* 97 (1): 298-317.
- Freeman, R. 1980. "The Exit-Voice Tradeoff in the Labor Market: Unionism, Job Tenure, Quits, and Separations." *Quarterly Journal of Economics* 94 (4): 643-673.
- Goldin, C. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091-1119.
- Gopal, R., Glasheen, J. J., Miyoshi, T. J., and Prochazka, A. V. 2005. "Burnout and Internal

- Medicine Resident Work-Hour Restrictions.” *JAMA Internal Medicine* 165 (22): 2595-2600.
- Gronau, R. 1988. “Sex-Related Wage Differentials and Women’s Interrupted Careers – the Chicken or the Egg.” *Journal of Labor Economics* 6 (3): 277-301.
- Heckman, J. J. and Willis, R. J. 1979. “The Distribution of Lifetime Labor Force Participation of Married Women: Reply to Mincer and Ofek.” *Journal of Political Economy*, 87 (1): 203-211.
- Hutter, M. M., Kellogg, K. C., Ferguson, C. M., Abbott, W. M., and Warshaw, A. L. 2006. “The Impact of the 80-Hour Resident Workweek on Surgical Residents and Attending Surgeons.” *Annals of Surgery* 243 (6): 864-875.
- Jones, A. M. and Jones, K. B. 2007. “The 88-Hour Family: Effects of the 80-Hour Work Week on Marriage and Childbirth in a Surgical Residency.” *Iowa Orthopaedic Journal* 27: 128-133.
- Karamanoukian, R. L., Ku, J. K., DeLaRosa, J., Karamanoukian, H. L., and Evans, G. R. 2006. “The Effects of Restricted Work Hours on Clinical Training.” *American Surgeon* 72 (1): 19-21.
- Kleven, H., Landais, C., and Søgaaard, J. E. 2018. “Children and Gender Inequality: Evidence from Denmark.” NBER Working Paper 24219.
- Kuo, Y.-F., Sharma, G., Freeman, J. L., and Goodwin, J. S. 2009. “Growth in the Care of Older Patients by Hospitalists in the United States.” *New England Journal of Medicine* 360 (11): 1102-1112.
- Landrigan, C. P., Barger, L. K., Cade, B. E., Ayas, N. T., and Czeisler, C. A. 2006. “Interns’ Compliance with Accreditation Council for Graduate Medical Education Work-Hour Limits.” *Journal of the American Medical Association* 296 (9): 1063-1070.
- Landrigan, C. P., Fahrenkopf, A. M., Lewin, D., Sharek, P. J., Barger, L. K., Eisner, M., Edwards, S., Chiang, V. W., Wiedermann, B. L., and Sectish, T. C. 2008. “Effects of the Accreditation Council for Graduate Medical Education Duty Hour Limits on Sleep, Work Hours, and Safety.” *Pediatrics* 122 (2): 250-258.
- Martini, S., Arfken, C. L., and Balon, R. 2006. “Comparison of Burnout among Medical Residents Before and After the Implementation of Work Hours Limits.” *Academic Psychiatry* 30 (4): 352-355.
- McBurney, P. G., Gustafson, K. K., and Darden, P. M. 2008. “Effect of 80-Hour Workweek on Continuity of Care.” *Clinical Pediatrics* 47 (8): 803-808.
- Melly, B. and Santangelo, G. 2015. “The Changes-in-Changes Model with Covariates.” Working paper.
- Nakayama, D. K., Thompson, W. M., Wynne, J. L., Dalton, M. L., Bozeman, A. T., and Innes, B. J. 2009. “The Effect of ACGME Duty Hour Restrictions on Operative Continuity of Care.” *American Surgeon* 75 (12): 1234-1237.

- Philibert, I., Friedmann, P., and Williams, W. T. 2002. "New Requirements for Resident Duty Hours." *Journal of the American Medical Association* 288 (9): 1112-1114.
- Philibert, I., Chang, B., Flynn, T., Friedmann, P., Minter, R., Scher, E., and Williams, W. T. 2009. "The 2003 Common Duty Hour Limits: Process, Outcome, and Lessons Learned." *Journal of Graduate Medical Education* 1 (2): 334-337.
- Philibert, I., Nasca, T., Brigham, T., and Shapiro, J. 2013. "Duty-Hour limits and Patient Care and Resident Outcomes: Can High-Quality Studies Offer Insight into Complex Relationships?" *Annual Review of Medicine* 64: 467-483.
- Rizzo, J. and Blumenthal, D. 1994. "Physician Labor Supply: Do Income Effects Matter?" *Journal of Health Economics* 13: 433-453.
- Rizzo, J. and Zeckhauser, R. 2007. "Pushing Incomes to Reference Points: Why Do Male Doctors Earn More?" *Journal of Economic Behavior and Organization* 63 (3): 514-536.
- Sarma, S., Thind, A., and Chu, M. K. 2011. "Do New Cohorts of Family Physicians Work Less Compared to Their Older Predecessors? The Evidence from Canada." *Social Science & Medicine* 71: 2049-2058.
- Sasser, A. 2005. "Gender Differences in Physician Pay: Tradeoffs between Career and Family." *Journal of Human Resources* 40 (2): 477-504.
- Scott, A. 2006. "The Productivity of the Health Workforce." *Australian Economic Review* 39 (3): 312-317.
- Staiger, D., Auerbach, D., and Buerhaus, P. 2009. "Comparison of Physician Workforce Estimates and Supply Projections." *Journal of the American Medical Association* 302 (15): 1674-1679.
- Staiger, D., Auerbach, D., and Buerhaus, P. 2010. "Trends in the Work Hours of Physicians in the United States." *Journal of the American Medical Association* 303 (8): 747-753.
- Szymczak, J. E., Brooks, J. V., Volpp, K. G., and Bosk, C. L. 2010. "To Leave or to Lie? Are Concerns about a Shift-Work Mentality and Eroding Professionalism as a Result of Duty-Hour Rules Justified?" *Milbank Quarterly* 88 (3): 350-381.
- Vidyarathi, A. R., Arora, V., Schnipper, J. L., Wall, S. D., and Wachter, R. M. 2006. "Managing Discontinuity in Academic Medical Centers: Strategies for a Safe and Effective Resident Sign-Out." *Journal of Hospital Medicine* 1 (4): 257-266.
- Wang, C. and Sweetman, A. 2013. "Gender, Family Status and Physician Labour Supply." *Social Science & Medicine* 94: 17-25.
- Wasserman, M. 2018. "Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents." Working paper.
- Watson, D., Slade S., Buske, L., and Tepper, J. 2006. "Intergenerational Differences in Workloads among Primary Care Physicians: A Ten-Year, Populations-Based Study." *Health*

Affairs 25 (6): 1620-1628.

Zhan, C., Miller, M. R., Wong, H., and Meyer, G. S. 2004 “The Effects of HMO Penetration on Preventable Hospitalizations.” *Health Services Research* 39 (2): 345-361.

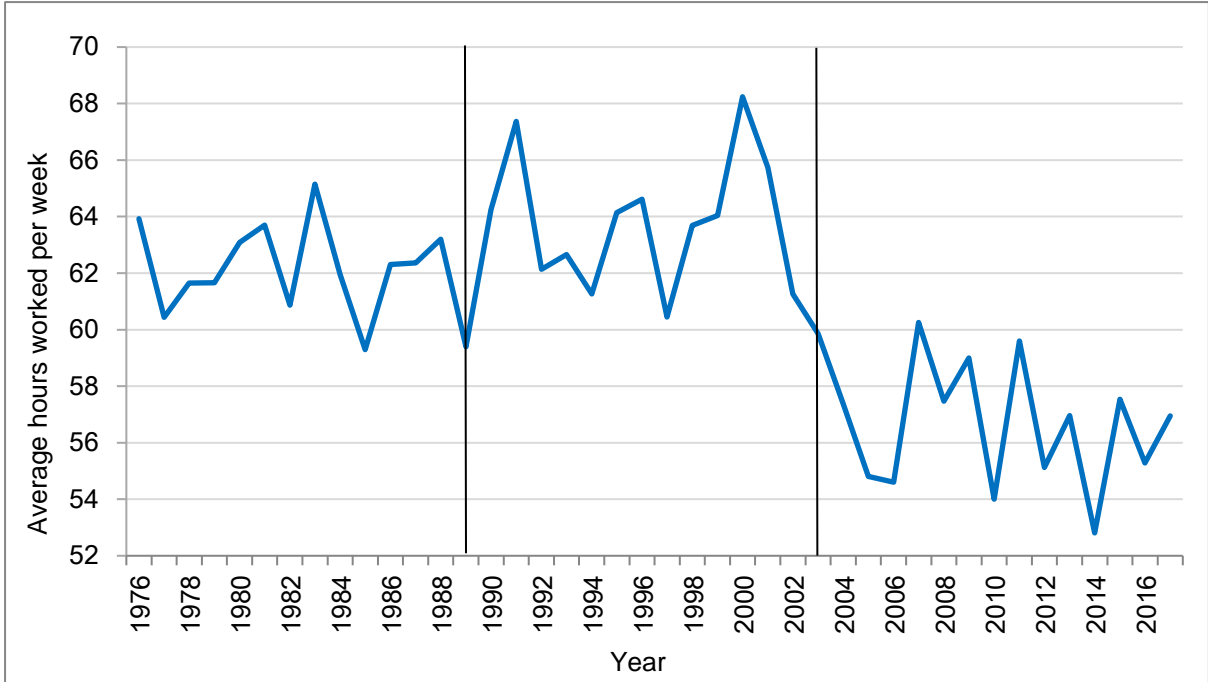


Figure 1: Average Hours Worked per Week by Residents, 1976-2017, Using the ASEC Data

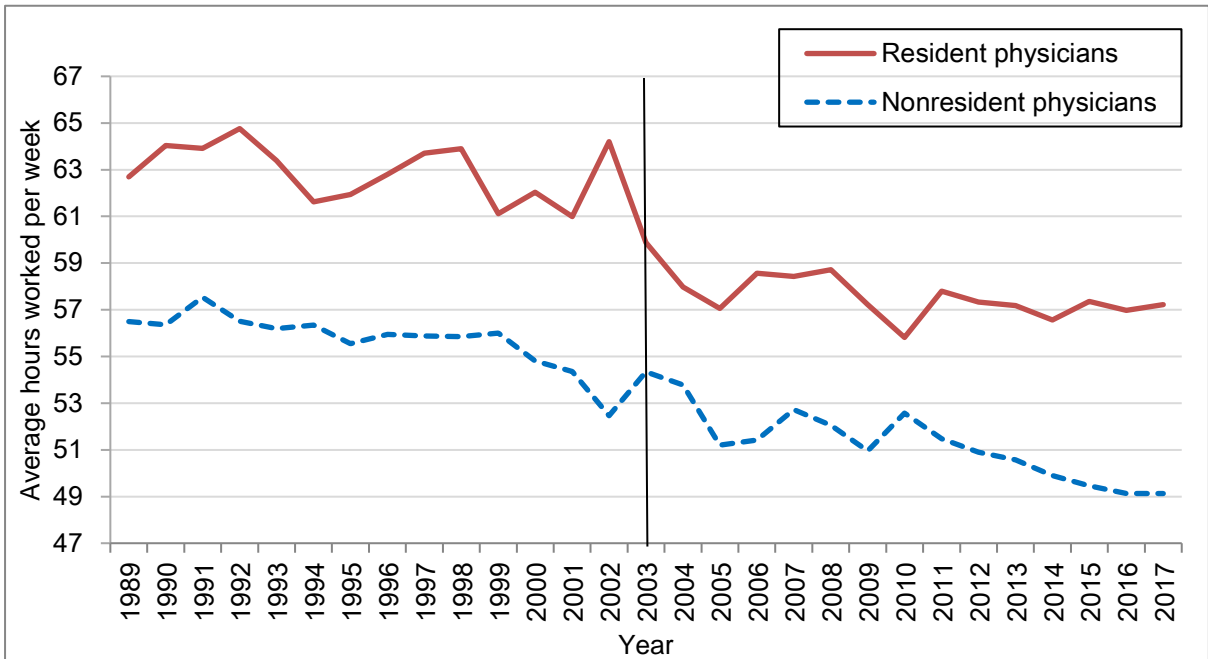


Figure 2: Average Hours Worked per Week, 1989-2017

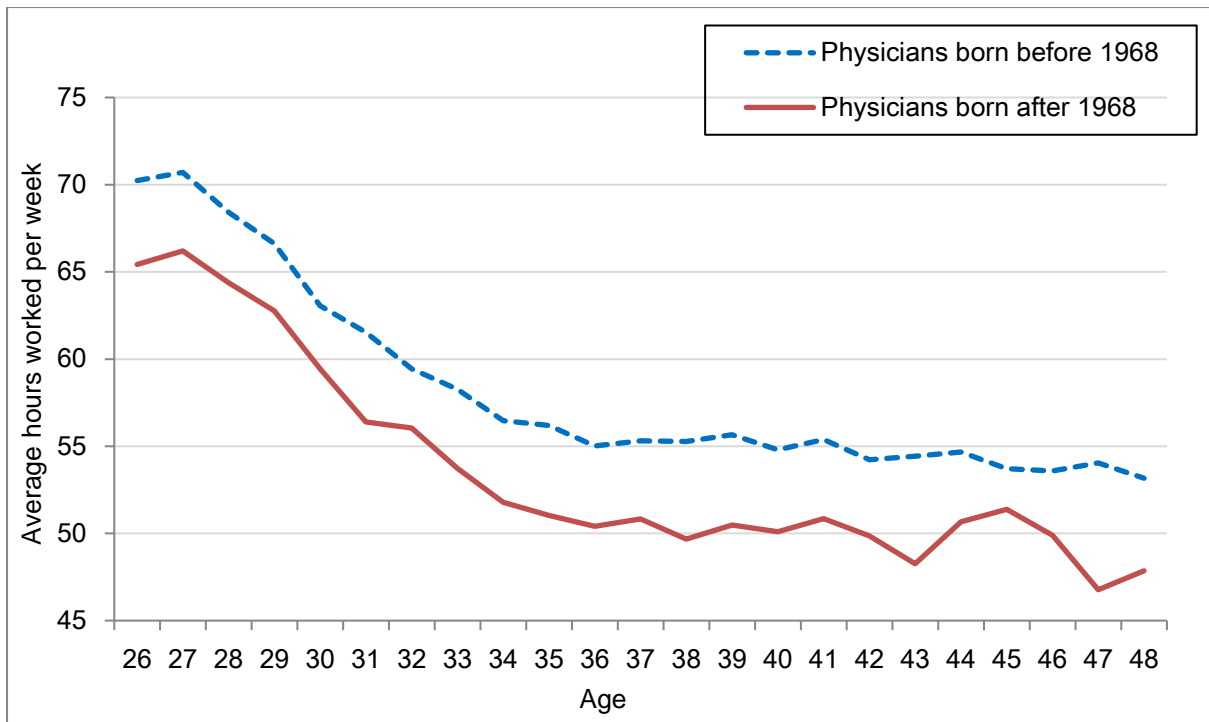


Figure 3: Age-Hours Profiles by Cohort

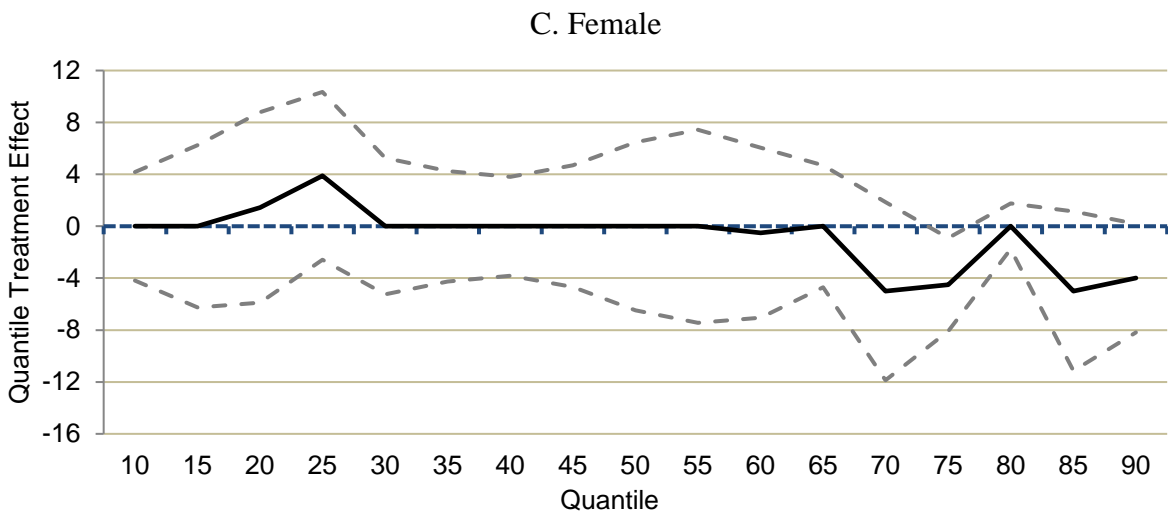
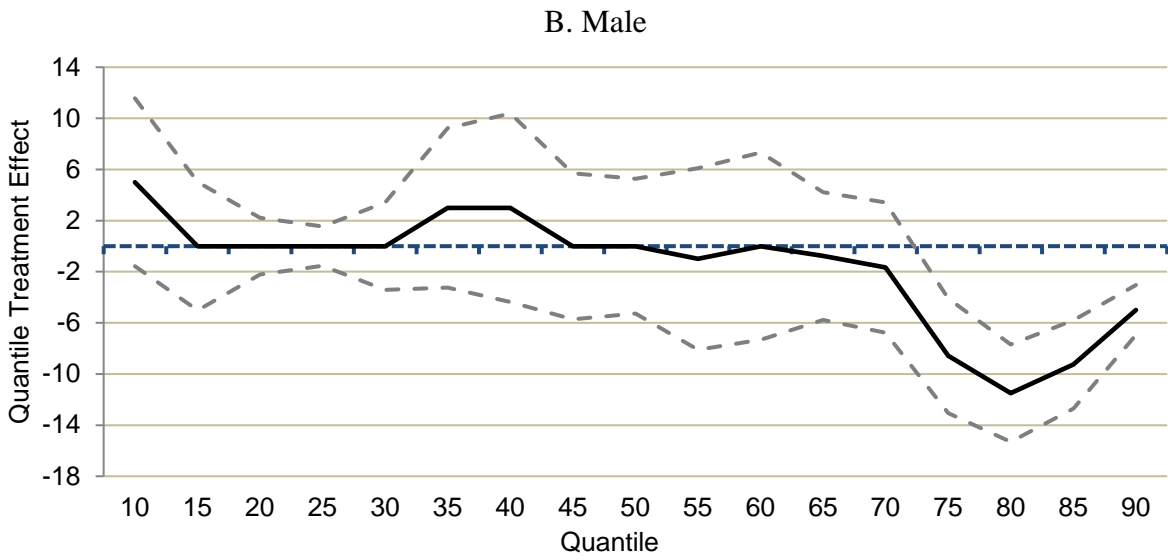
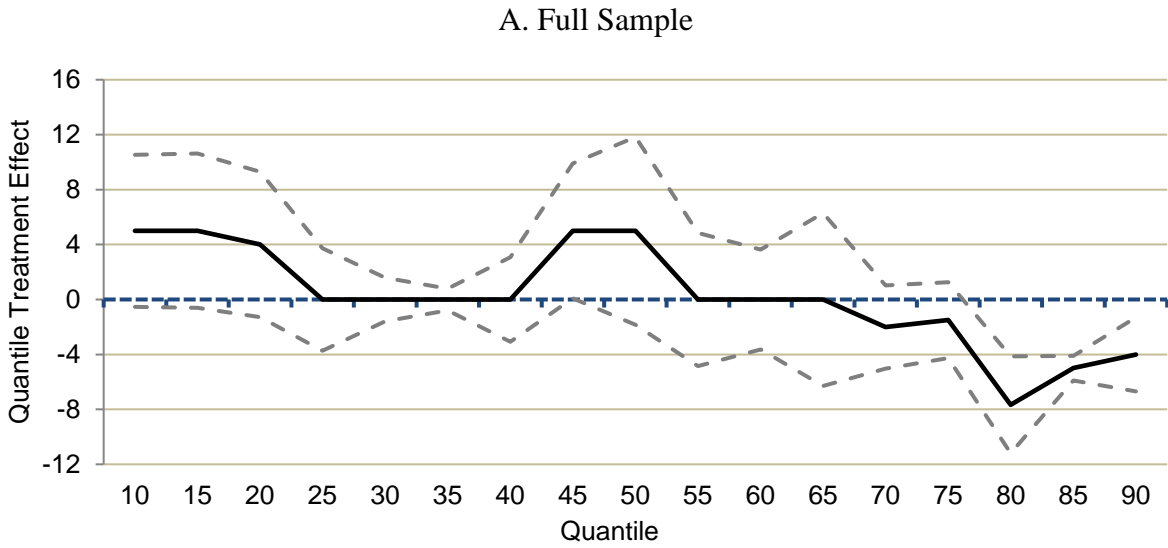


Figure 4: Changes-in-Changes Estimates

Notes: These figures show the changes-in-changes estimates of the reform on long-term labor supply for seventeen quantile values, controlling for age, gender, and race/ethnicity. Dotted lines provide bootstrapped pointwise 95-percent confidence intervals for quantile treatment effects with 1,000 replications.

Table 1: Summary Statistics

	All	Gender		Treatment Status	
		Male	Female	Treatment	Control
Age	37.586 (6.234)	38.075 (6.209)	36.614 (6.170)	34.164 (5.255)	40.205 (5.616)
Male	0.665 (0.472)			0.579 (0.494)	0.736 (0.441)
White	0.696 (0.460)	0.720 (0.449)	0.649 (0.477)	0.622 (0.485)	0.756 (0.430)
Black	0.059 (0.236)	0.047 (0.211)	0.084 (0.277)	0.069 (0.253)	0.051 (0.219)
Hispanic	0.057 (0.231)	0.058 (0.233)	0.055 (0.228)	0.063 (0.244)	0.054 (0.226)
Asian	0.182 (0.386)	0.170 (0.376)	0.206 (0.404)	0.240 (0.427)	0.134 (0.341)
Other race	0.006 (0.077)	0.005 (0.073)	0.007 (0.084)	0.006 (0.079)	0.006 (0.075)
Married	0.755 (0.430)	0.783 (0.412)	0.700 (0.458)	0.692 (0.462)	0.806 (0.396)
Number of children	1.312 (1.290)	1.439 (1.336)	1.060 (1.154)	0.971 (1.181)	1.575 (1.313)
Family Size	3.186 (1.566)	3.334 (1.599)	2.892 (1.454)	2.810 (1.478)	3.472 (1.572)
Weekly hours worked	55.536 (17.376)	57.614 (16.831)	51.410 (17.701)	55.271 (17.600)	55.813 (17.203)
Exposure	0.438 (0.496)	0.381 (0.486)	0.554 (0.497)		
Observations	69,371	46,160	23,211	26,516	40,615

Notes: Table reports means and standard deviations (in parentheses), weighted using sampling weights. The sample includes active physicians aged 26-48 in the monthly CPS between 1989 and 2017. The treatment group includes physicians born after 1968, and the control group includes physicians born before 1968.

Table 2: The Impact of the Reform on Resident Hours

Outcome: Hours worked at ages 26-34	All	Male	Female
Exposed	-8.399* (4.601)	-10.538** (5.249)	-5.939 (7.706)
Age	-1.989*** (0.110)	-1.821*** (0.137)	-2.251*** (0.186)
Male	3.941*** (0.502)		
Black	-1.334 (1.123)	-2.888* (1.715)	-0.038 (1.449)
Hispanic	0.335 (1.082)	0.371 (1.329)	0.903 (1.752)
Asian	-2.934*** (0.603)	-2.896*** (0.777)	-2.815*** (0.943)
Other race	4.861 (3.357)	5.550 (5.366)	3.276 (3.460)
Cohort FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
State FEs	Yes	Yes	Yes
Observations	22,347	13,473	8,874

Notes: The dependent variable is weekly hours worked at ages 26 to 34. Standard errors are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights and adjusted for respondent-level clustering.

Table 3: Difference-in-differences Estimates of the Impact on Long-Term Labor Supply

	Ages 26-34 (Residency)	Ages 35-48 (Post-Residency)	Difference (%)
<i>A. Full Sample</i>			
Treatment group	59.198	50.284	-8.914 (-15.058%)
Control group	61.477	54.545	-6.932 (-11.276%)
Difference	-2.279	-4.261	
Difference-in-differences [Standard Error]			-1.982 (-3.782%) [0.660]***
<i>B. Male</i>			
Treatment group	60.639	53.125	-7.514 (-12.391%)
Control group	63.168	56.759	-6.409 (-10.146%)
Difference	-2.529	-3.634	
Difference-in-differences [Standard Error]			-1.106 (-2.245%) [0.801]
<i>C. Female</i>			
Treatment group	57.290	46.167	-11.123 (-19.415%)
Control group	57.415	48.155	-9.260 (-16.128%)
Difference	-0.125	-1.988	
Difference-in-differences [Standard Error]			-1.863 (-3.287%) [1.133]*

Note: The sample includes active physicians aged 26-48 in the monthly CPS between 1989 and 2017. The treatment group includes physicians born after 1968, and the control group includes physicians born before 1968. Data are weighted using CPS sampling weights, and the standard errors are adjusted for respondent-level clustering. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

Table 4: Estimation Results of the Impact on Long-Term Labor Supply

Outcome: Hours worked	All		Male		Female	
	(1)	(2)	(3)	(4)	(5)	(6)
Exposed*Post	-4.326*** (0.657)	-4.226*** (0.638)	-3.962*** (0.843)	-3.883*** (0.837)	-4.457*** (0.979)	-4.431*** (0.977)
Age		-0.608*** (0.041)		-0.548*** (0.051)		-0.714*** (0.074)
Male		6.339*** (0.304)				
Black		-0.387 (0.711)		-0.742 (1.017)		-0.037 (0.966)
Hispanic		-0.332 (0.654)		-0.619 (0.775)		0.407 (1.173)
Asian		-2.576*** (0.362)		-2.863*** (0.442)		-2.133*** (0.615)
Other race		3.903** (1.792)		4.496* (2.482)		3.124 (2.316)
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	No	Yes	No	Yes	No	Yes
Observations	67,131	67,131	44,851	44,851	22,280	22,280

Notes: The sample includes active physicians aged 26-48 in the monthly CPS between 1989 and 2017. Standard errors are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights and adjusted for respondent-level clustering.

Table 5: The Impact of the Reform on Marriage and Divorce

	All		Male		Female	
	Married (1)	Divorced (2)	Married (3)	Divorced (4)	Married (5)	Divorced (6)
Exposed	0.049*** (0.015)	-0.001 (0.007)	0.027 (0.019)	0.007 (0.008)	0.092*** (0.028)	-0.014 (0.014)
Age	0.019*** (0.001)	0.004*** (0.0004)	0.019*** (0.001)	0.003*** (0.0004)	0.018*** (0.002)	0.005*** (0.001)
Male	0.044*** (0.009)	-0.027*** (0.004)				
Black	-0.148*** (0.021)	-0.001 (0.008)	-0.133*** (0.027)	0.003 (0.011)	-0.163*** (0.032)	-0.005 (0.012)
Hispanic	-0.007 (0.019)	0.010 (0.010)	-0.006 (0.023)	0.019 (0.012)	-0.007 (0.035)	-0.007 (0.014)
Asian	0.040*** (0.011)	-0.027*** (0.004)	0.020 (0.014)	-0.022*** (0.005)	0.072*** (0.019)	-0.034*** (0.007)
Other race	-0.069 (0.051)	-0.016 (0.011)	0.011 (0.058)	-0.010 (0.015)	-0.191** (0.083)	-0.031* (0.018)
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67,131	67,131	44,851	44,851	22,280	22,280

Notes: The sample includes active physicians aged 26-48 in the monthly CPS between 1989 and 2017. The dependent variable are dummy variables for being married or divorced on the timing of the survey. Standard errors are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights and adjusted for respondent-level clustering.

Table 6: The Impact of the Reform on Fertility Decisions

	All		Male		Female	
	Extensive	Intensive	Extensive	Intensive	Extensive	Intensive
	(1)	(2)	(3)	(4)	(5)	(6)
Exposed	0.025 (0.016)	0.037 (0.042)	-0.004 (0.020)	-0.049 (0.054)	0.090*** (0.027)	0.202*** (0.064)
Age	0.034*** (0.001)	0.092*** (0.002)	0.032*** (0.001)	0.091*** (0.003)	0.039*** (0.001)	0.092*** (0.003)
Male	0.040*** (0.010)	0.213*** (0.024)				
Black	-0.105*** (0.021)	-0.243*** (0.055)	-0.151*** (0.029)	-0.333*** (0.081)	-0.046 (0.030)	-0.140* (0.072)
Hispanic	-0.030 (0.020)	-0.109** (0.052)	-0.015 (0.025)	-0.063 (0.068)	-0.059* (0.034)	-0.188** (0.075)
Asian	0.017 (0.012)	-0.076*** (0.028)	0.005 (0.015)	-0.117*** (0.037)	0.040** (0.019)	-0.008 (0.043)
Other race	-0.048 (0.050)	-0.170 (0.138)	-0.005 (0.068)	-0.063 (0.206)	-0.109* (0.065)	-0.336** (0.145)
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	67,131	67,131	44,851	44,851	22,280	22,280

Notes: The sample includes active physicians aged 26-48 in the monthly CPS between 1989 and 2017. The dependent variables are: (1) fertility at the extensive margin: an indicator of having at least one child and (2) fertility at the intensive margin: the number of children on the time of the survey. Standard errors are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights and adjusted for respondent-level clustering.

Appendix

Table A1: Estimation Results of the Impact on Long-Term Labor Supply
– Using Tighter Year-of-Birth Windows for Treatment and Control Groups

Treatment	Control	Eq. (3), no controls	Eq. (3), controls	Observations
<i>All</i>				
After 1968	Before 1968	-4.326***	-4.226***	67,131
	(Main analysis)	(0.657)	(0.638)	
After 1969	Before 1967	-4.856***	-4.849***	56,658
		(0.877)	(0.849)	
After 1970	Before 1966	-4.977***	-4.947***	47,088
		(1.255)	(1.223)	
After 1971	Before 1965	-4.548***	-4.342***	38,642
		(1.664)	(1.637)	
<i>Male</i>				
After 1968	Before 1968	-3.962***	-3.883***	44,851
	(Main analysis)	(0.843)	(0.837)	
After 1969	Before 1967	-4.470***	-4.369***	38,253
		(1.089)	(1.085)	
After 1970	Before 1966	-4.849***	-4.737***	32,065
		(1.547)	(1.545)	
After 1971	Before 1965	-3.944**	-3.622*	26,716
		(1.937)	(1.929)	
<i>Female</i>				
After 1968	Before 1968	-4.457***	-4.431***	22,280
	(Main analysis)	(0.979)	(0.977)	
After 1969	Before 1967	-5.097***	-5.028***	18,405
		(1.352)	(1.346)	
After 1970	Before 1966	-4.761**	-4.711**	15,023
		(1.989)	(1.974)	
After 1971	Before 1965	-5.142*	-5.085*	11,926
		(2.857)	(2.834)	

Notes: Each cell contains an estimate of the effect on post-residency hours. Standard errors are in parentheses. Data are weighted using CPS sampling weights and adjusted for respondent-level clustering. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. Three alternative specifications are used to test the robustness of the results. The treatment status (defined by year of birth) and residency status (defined by age) associated with each row are as follows, where the first one is used in the main analysis.

- (1) Treatment: born after 1968; Control: born before 1968. Residency: ages 26-34; Post-residency: ages 35-48.
- (2) Treatment: born after 1969; Control: born before 1967. Residency: ages 26-33; Post-residency: ages 36-48.
- (3) Treatment: born after 1970; Control: born before 1966. Residency: ages 26-32; Post-residency: ages 37-48.
- (4) Treatment: born after 1971; Control: born before 1965. Residency: ages 26-31; Post-residency: ages 38-48.

Table A2: Estimation Results of the Impact on Long-Term Labor Supply
– Excluding Nonrandom Selection into Treatment

	Eq. (3), no controls	Eq. (3), controls	Observations
<i>All</i>			
1941-1991 Cohorts (Main analysis)	-4.326*** (0.657)	-4.226*** (0.638)	67,131
1941-1980 Cohorts	-5.008*** (0.704)	-4.894*** (0.680)	61,481
1941-1978 Cohorts	-5.455*** (0.758)	-5.394*** (0.732)	59,169
1941-1976 Cohorts	-5.810*** (0.820)	-5.769*** (0.789)	56,259
<i>Male</i>			
1941-1991 Cohorts (Main analysis)	-3.962*** (0.843)	-3.883*** (0.837)	44,851
1941-1980 Cohorts	-4.263*** (0.888)	-4.146*** (0.881)	41,899
1941-1978 Cohorts	-4.668*** (0.947)	-4.521*** (0.940)	40,714
1941-1976 Cohorts	-4.586*** (1.020)	-4.475*** (1.011)	39,089
<i>Female</i>			
1941-1991 Cohorts (Main analysis)	-4.457*** (0.979)	-4.431*** (0.977)	22,280
1941-1980 Cohorts	-5.759*** (1.066)	-5.733*** (1.063)	19,582
1941-1978 Cohorts	-6.384*** (1.156)	-6.397*** (1.153)	18,455
1941-1976 Cohorts	-7.414*** (1.254)	-7.412*** (1.249)	17,170

Notes: Each cell contains an estimate of the effect on hours worked after residency. Standard errors are in parentheses. Data are weighted using CPS sampling weights and adjusted for respondent-level clustering. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

Table A3: Percentiles of the Distribution of Weekly Hours Worked

	Observations	Mean (SD)	Percentiles of the distribution						
			5	10	25	50	75	90	95
<i>All, ages 26-34</i>									
Treatment group	14,677	59.198 (18.166)	35	40	40	60	72	80	90
Control group	7,670	61.477 (19.370)	35	40	49	60	75	90	99
<i>All, ages 35-48</i>									
Treatment group	11,839	50.284 (15.480)	30	35	40	50	60	70	80
Control group	32,945	54.545 (16.415)	30	40	40	50	60	80	84
<i>Male, ages 26-34</i>									
Treatment group	8,127	60.639 (17.831)	40	40	45	60	75	80	95
Control group	5,346	63.168 (18.673)	40	40	50	60	80	95	99
<i>Male, ages 35-48</i>									
Treatment group	6,906	53.125 (15.272)	35	40	40	50	60	75	80
Control group	24,472	56.759 (15.935)	36	40	45	55	65	80	85
<i>Female, ages 26-34</i>									
Treatment group	6,550	57.290 (18.429)	32	40	40	55	70	80	90
Control group	2,324	57.415 (20.389)	29	35	40	55	70	90	99
<i>Female, ages 35-48</i>									
Treatment group	4,933	46.167 (14.841)	24	30	40	40	55	65	78
Control group	8,473	48.155 (16.102)	24	30	40	45	60	70	80

Notes: Table reports means, standard deviations (in parentheses), and percentiles of weekly hours worked, weighted using sampling weights. The sample includes active physicians aged 26-48 in the monthly CPS between 1989 and 2017. The treatment group includes physicians born after 1968, and the control group includes physicians born before 1968.

Table A4: The Impact of the Reform on Marriage and Fertility Decisions

		All	Male	Female
Married	Probit	0.048*** (0.015)	0.028 (0.018)	0.089*** (0.028)
	Logit	0.049*** (0.016)	0.028 (0.018)	0.089*** (0.028)
Divorced	Probit	-0.001 (0.007)	0.007 (0.008)	-0.012 (0.012)
	Logit	-0.001 (0.007)	0.006 (0.009)	-0.012 (0.013)
Fertility (Extensive)	Probit	0.025 (0.016)	0.001 (0.020)	0.079*** (0.028)
	Logit	0.027* (0.016)	0.005 (0.020)	0.080*** (0.029)

Notes: The sample includes active physicians aged 26-48 in the monthly CPS between 1989 and 2017. The independent variables include exposure to the regulations, age, gender, race/ethnicity, year fixed effects, and state fixed effects. Standard errors are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights and adjusted for respondent-level clustering. The effects is based on the average marginal effects using probit and logit models.

Table A5: The Impact of the Reform on Marriage and Fertility Decisions

		All	Male	Female
Married	Ages 26-34	0.011 (0.039)	-0.033 (0.049)	0.072 (0.062)
	Ages 35-48	0.041** (0.019)	0.021 (0.022)	0.083** (0.037)
Divorced	Ages 26-34	0.005 (0.013)	0.015 (0.021)	-0.007 (0.010)
	Ages 35-48	0.001 (0.011)	0.007 (0.011)	-0.008 (0.023)
Fertility (Extensive)	Ages 26-34	-0.042 (0.035)	-0.075* (0.045)	0.009 (0.055)
	Ages 35-48	0.019 (0.022)	-0.002 (0.027)	0.066* (0.038)
Fertility (Intensive)	Ages 26-34	-0.067 (0.073)	-0.118 (0.099)	0.019 (0.092)
	Ages 35-48	-0.010 (0.065)	-0.084 (0.084)	0.131 (0.099)

Notes: The sample includes active physicians aged 26-48 in the monthly CPS between 1989 and 2017. The independent variables include exposure to the regulations, age, gender, race/ethnicity, year fixed effects, and state fixed effects. Standard errors are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights and adjusted for respondent-level clustering.