Table 2: Estimated average smoking effects on infant birth weight

	Estimate (g)
Evans and Ringel (1999)	-600 − -360
Almond et al. (2005)	-203.2
Abrevaya (2006)	-144 − -178
Arellano and Bonhomme (2011)	-161

Most existing studies used the Natality Data by the National Center for Health Statistics (NCHS) for its large sample size and a wealth of information on covariates. The birth data is based on birth records from every live birth in the U.S. and contains detailed information on birth outcomes, maternal prenatal behavior and medical status, and demographic attributes.[14] Table 1 describes the data used in the recent literature.

While some studies such as Hoderlein and Sasaki (2011) and Caetano (2012) use the number of cigarettes per day as a continuous treatment variable, most applied research uses a binary variable for smoking. The literature, including Evans and Farrelly (1998), found that individuals, especially women, tend to underreport their cigarette consumption. On the other hand, smoking participation has shown to be more accurately reported among adults in the literature. Moreover, the literature has pointed out that the number of cigarettes may not be a good proxy for the level of nicotine intake. Previous studies, including Chaloupka and Warner (2000), Evans and Farrelly (1998), Farrelly et al. (2004), Adda and Cornaglia (2006), and Abrevaya and Puzzello (2012) discussed that although an increase in cigarette taxes leads to a lower percentage of smokers and less cigarettes consumed by smokers, it causes individuals to purchase cigarettes that contain more tar and nicotine as compensatory behavior.

Although many recent studies are based on the same NCHS data set, their estimates of average smoking effects are quite varied, ranging from -144 grams to -600 grams depending on their estimation methods and samples. Table 2 summarizes their estimates.

5.3 Data

I use the NCHS Natality dataset. My sample consists of births to women who were in their first trimester during the period between two years before and two years after the tax increase. In other words, I consider births to women who conceived babies in MA between October 1990 and September 1994.[15] I define the instrument as an indicator of whether the agent faces the high tax rate from the tax hike during the first trimester of pregnancy. Since the tax increase occurred in MA in January of 1993, the instrument Z can be

[14] Unfortunately the Natality Data does not provide information on mothers' income and weight.

[15] To trace the month of conception, I use information on the month of birth and the clinical estimate of gestation weeks.