



Analysis of Women's Labor-Force Participation Dataset

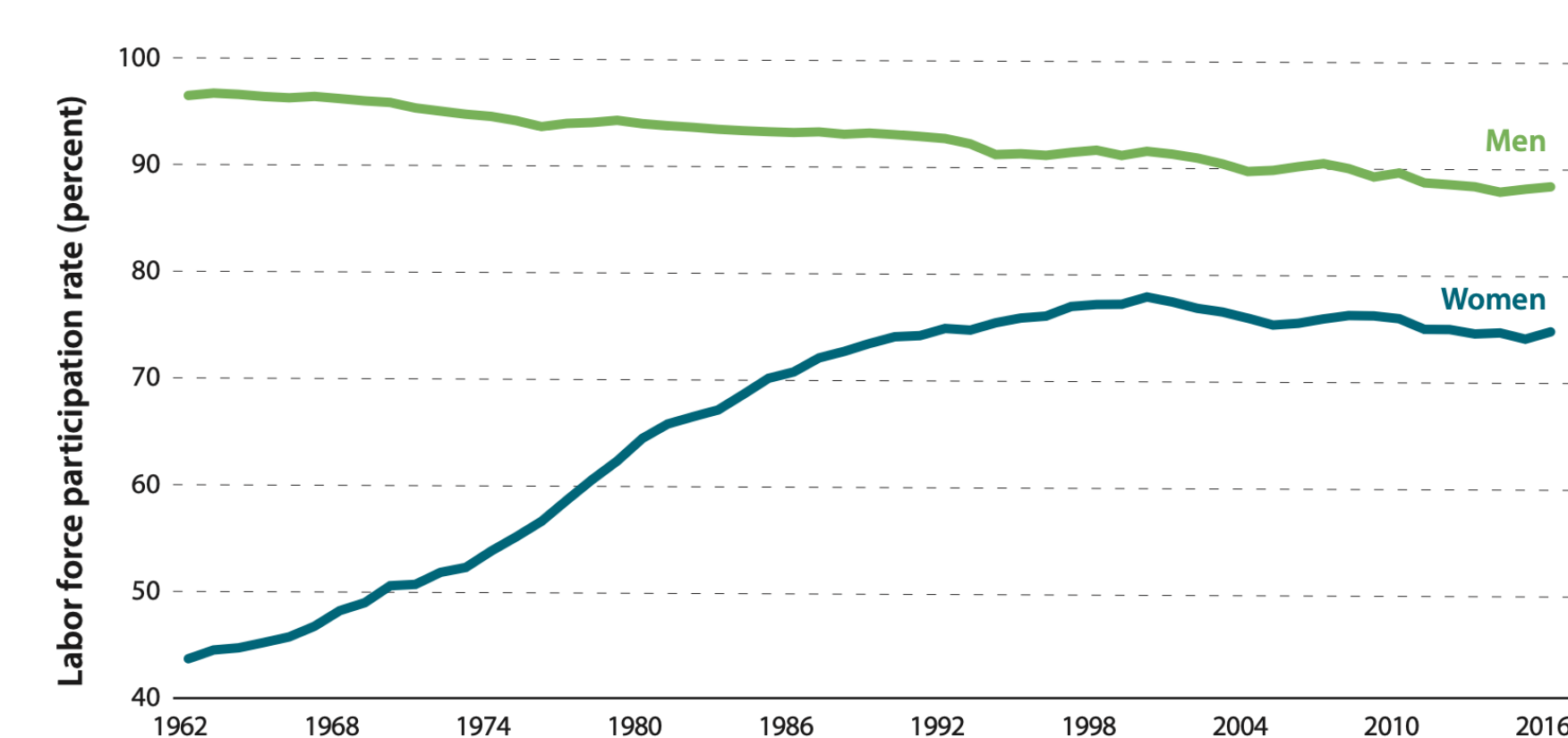
Han Chen

Department of Statistics, University of California, Davis

Background

In almost every country in the world, men are more likely to participate in labor markets than women. However, as shown in figure 1, these gender differences in participation rates have been narrowing substantially. This trend has driven us to find out the potential predictors that is significant in women's labor force participation

Prime-Age Labor Force Participation, by Gender



Source: Current Population Survey Annual Social and Economic Supplement 1962-2016.
Note: Prime age indicates ages 25 to 54.

Figure 1: Women's Labor-Force Participation by Gender

Data Description

The data sample consists of 753 married white women between the ages of 30 and 60 in 1975, with 428 working at some time during the year.

- Y : Response, Binary, Women's labor-force participation
- X_1 : Discrete, Number of children whose ages are less than five
- X_2 : Discrete, Number of children whose ages between six to nineteen
- X_3 : Discrete, Age of women
- X_4 : Binary, Wife's college attendance
- X_5 : Binary, Husband's college attendance
- X_6 : Continuous, Log expected wage rate
- X_7 : Continuous, Family income exclusive wife's income

Smoothing Method and GAM

We use Generalized Additive model to detect some patterns of predictors.

- There is a huge difference in X_6 for glm and gam. The quadratic form means that for woman who is poor and very rich, they are more likely to work.
- There is not huge difference in X_7 , which means that it is not very necessary to use smoothing term for X_7
- There exist some undersmoothing in figure 2(b)

Visualization of Smoothing

We use smoothing methods to detect the relationship between X_6 and Y , and the relationship between X_7 and Y in figure 2

- Detect a U-shape relationship between X_6 and Y
- X_7 and Y are almost linear relationship.

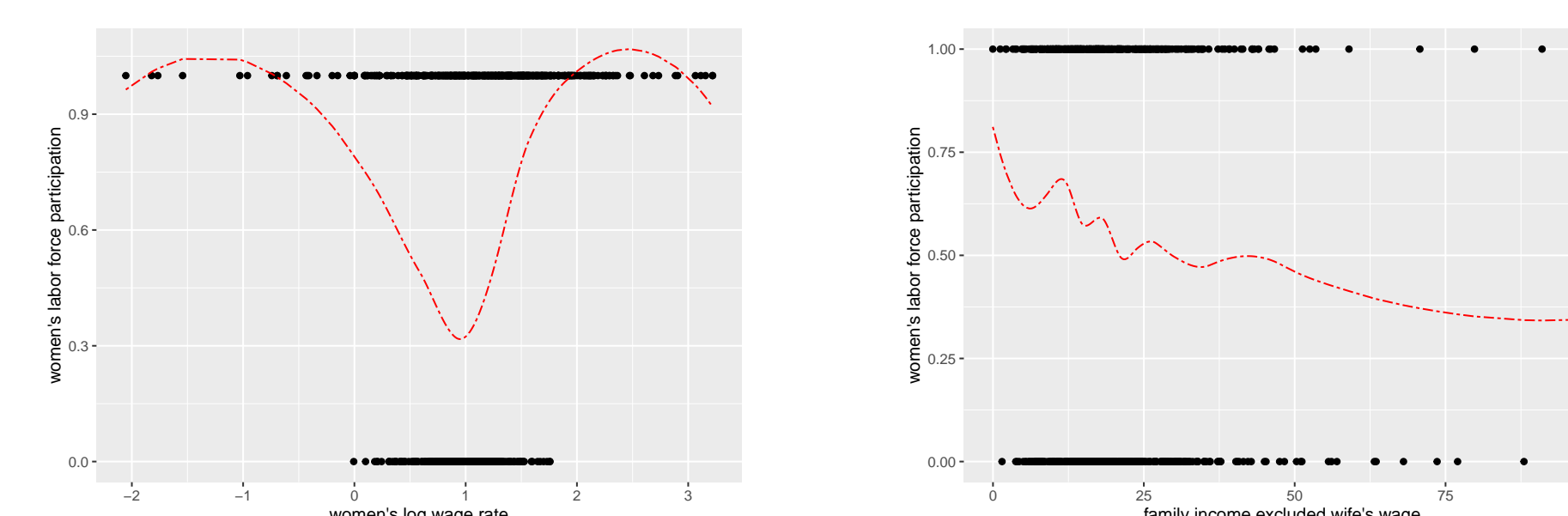


Figure 2:(a)Smoothing between X_6 and Y (b)Smooth between X_7 and Y

Main Result

From the smoothing part and GAM model part, we decide to only consider predictors, X_1 , X_3 , X_6 , X_6^2 , and X_7 . We use the cross validation to select the final model.

$$\text{Logit}(\mathbb{E}(Y|X)) = 6.547 - 1.516X_1 - 0.068X_3 - 7.850X_6 + 4.648X_6^2 - 0.032X_7$$

With Y follows binomial regression. We show the anova parametric table in table 1 and residual plots in figure 3

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	6.5500	0.82100	7.97	1.55e-15
k5	-1.5200	0.22200	-6.83	8.61e-12
age	-0.0686	0.01280	-5.37	7.99e-08
lwg	-7.8500	1.09000	-7.22	5.09e-13
I(lwg ²)	4.6500	0.55600	8.36	<2e-16
inc	-0.0325	0.00852	-3.81	1.39e-04

Table 1:Parametric Anova of GLM model

- We fix the previous lack of fit problem, the smoothing figure is much flat especially in the tail
- In a recent study of lundborg2017can, show that women are successfully treated by IVF(in vitro fertilization) in Denmark are less likely to attend labor force market because of having children.
- For low-income women, where the agricultural work may be particularly important for them. Under such circumstances, family work and productive agricultural work can be more easily reconciled
- For high income family, women have more option of moving into paid and white-collar work
- For those middle-income women, their target work are more capital intensive and physically separated from home thus contributing to a decrease in women's labor force participation.

Discussion

Our further analysis will focus on the analysis of the trend in figure 1. Whether or not these key factors changed over time. And if there is some other new factors included.

Conclusion

In this article, we first use GLM method to fit the model and do the model diagnosis to detect some trend of underfitting. And then we turn to smoothing curve and GAM model and modify our model. We detect a U-Shape between women's log wage and their force participation, which can be explained for different income level family. In addition, lower rates of fertility and younger age can also free up women's time to join the market.

Acknowledgements

Thank you to Professor Hans Mueller's guidance. Thank you to Dr. Yaqing Chen's help and lab.

References

- [1] John Ashworth Nelder and Robert WM Wedderburn. Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, 135(3):370-384, 1972.
- [2] Thomas A Mroz. The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. *Econometrica: Journal of the Econometric Society*, pages 765-799, 1987.
- [3] Petter Lundborg, Erik Plug, and Astrid Würtz Rasmussen. Can women have children and a career? iv evidence from ivf treatments. *American Economic Review*, 107(6):1611-37, 2017.
- [4] Joan Entmacher, Lauren Frohlich, Katherine Gallagher Robbins, Emily Martin, and Liz Watson. Underpaid and overloaded: Women in low-wage jobs. *Washington, DC: National Women's Law Center*, 2014.

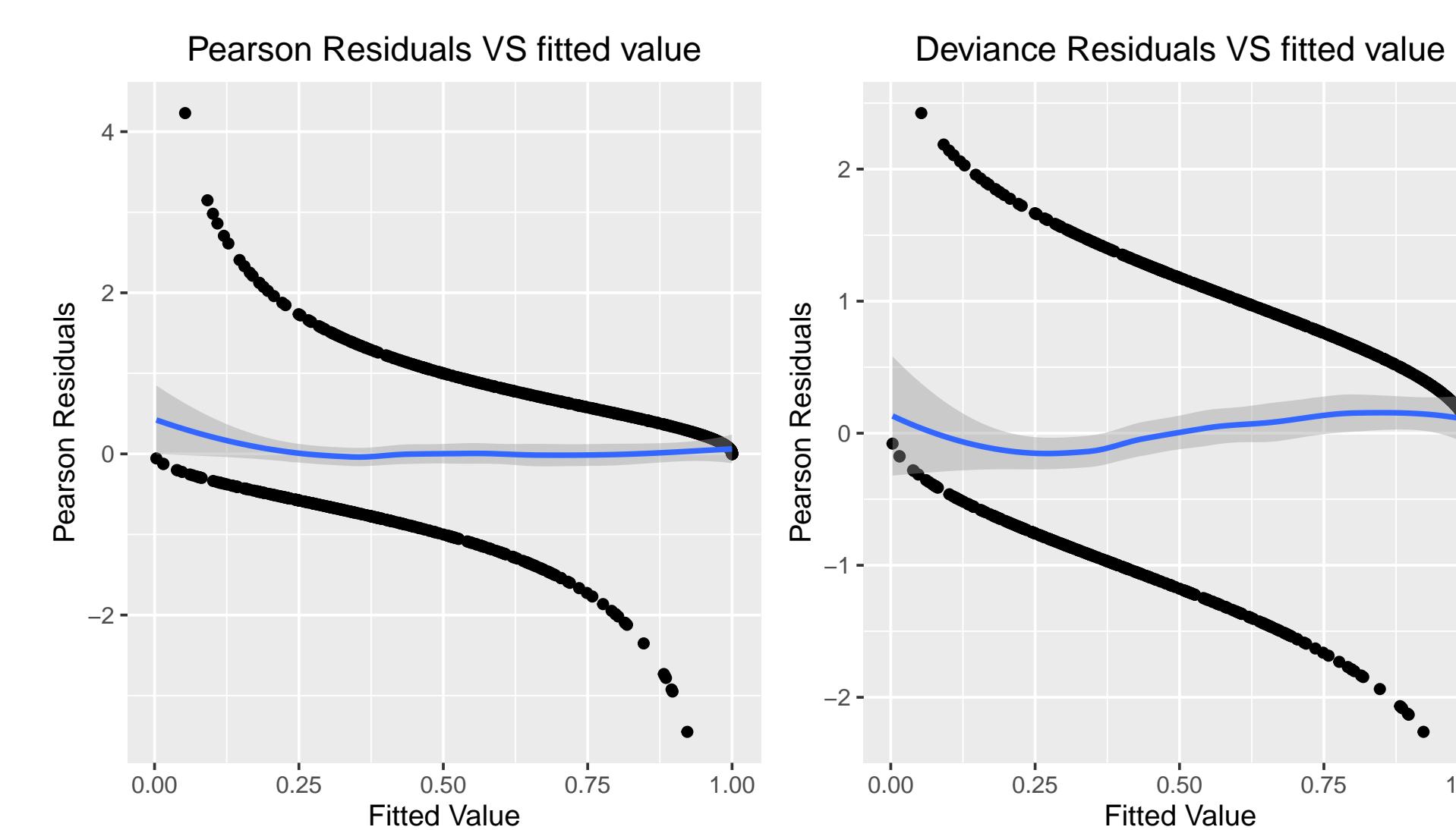


Figure 3:Residuals VS fitted value