Medicare Payments and Provider Analysis

Adam Wells

Dec. 18, 2019

Overview

According to the Centers for Medicare and Medicaid Service (CMS), Healthcare costs are equivalent to about 18% of the United States' GDP, and roughly 20% of all healthcare spending is related to Medicare and Medicaid. In order to better understand the distribution of medicare payments to healthcare providers, this report uses exploratory data techniques to analyze data sets known as the "Medicare Fee-for-Service Provider Utilization and Payment Data, Physician and Other Supplier Public Use File."

Some highlights from my analysis:

- The distribution of payments to providers was skewed, with top providers receiving an out-sized share of payments. The top 10% of individual providers (ranked by total payments) received 56.5% of payments, and the top 50% of providers received 94.9% of payments. Payments to organizational providers were even more concentrated: The top 10% of organizational providers received 86.7% of payments, and the top 50% of providers received 99.2% of payments.
- The highest paid organizational provider, a clinical laboratory in NC, collected $873.86 million between 2012 and 2016.
- The highest paid individual provider, an ophthalmologist in FL, collected $52.72 million between 2012 and 2016. This provider performed 2.98299×10^5 services during that time (approximately 270 services per day!).
- Concentration in payments reflected concentration in services: The top 10% of individual providers received 56.5% of payments and performed 60.68% of services. The top 10% of organizational providers received 86.7% of payments and performed 88.82% of services.
- There was a significant gender disparity in average payments made to individual providers. Women received $11.9 (95% confidence interval: (-17.02, -6.79) less per service on average than their male counterparts with the same "provider type" designation–a 13.14% difference.

Data Overview

This report examines the "Medicare Fee-for-Service Provider Utilization and Payment Data, Physician and Other Supplier Public Use File," which was compiled by the Office of Information Products and Data Analytics of the Centers for Medicare and Medicaid Services. The data sets contain "information on utilization, payment (allowed amount and Medicare payment), and submitted charges by National Provider Identifier (NPI), type of provider, Healthcare Common Procedure Coding System (HCPCS) code, and place of service" from 2012-2017. Each observation in the data sets corresponds to a particular NPI, which allows for easy cross-referencing between data sets. The data sets contain information on the provider, including credential and geographical location (zip code, state), the service provided (medicare indicator, service place, hcpcs code), the amount paid (amount charged, amount allowed, amount paid), and so forth.

These data sets is very large– approximately 56 million observations and 19 variables. Manipulating large data sets is challenging. R holds data in the computer's memory (RAM), and my computer's memory (8gb) was quickly exhausted. I therefore used the ff package to create a virtual database, which is only partially held in memory. From that virtual database, I was able to create new, smaller databases when necessary. I also made two main modifications to the data. First, I added a column of total payments (avg.payment * svc.count) in order to gain a better understanding of payments made to providers. Second, I joined the "claims" data frame to the "providers" data frame in order to explore characteristics of the providers (e.g., gender, credentials, and so forth.)

Research Questions & Exploratory Data Analysis

In this section, I will use exploratory data techniques to analyze the overall "shape" of the data and to generate topics for further research. The following charts summarize payment data for individual (IP) and organizational (OP) providers:

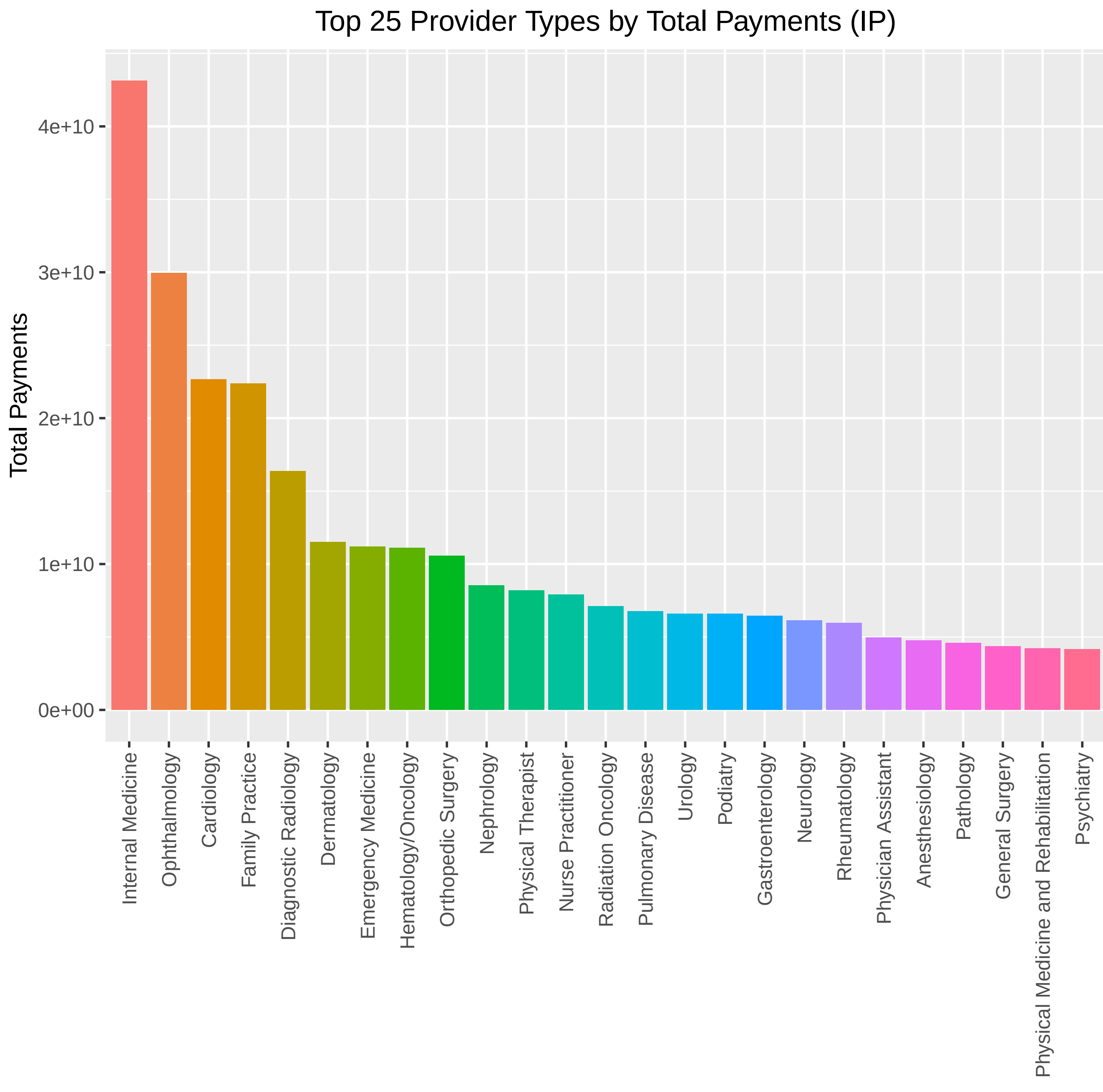
Table 1: Payments per Service

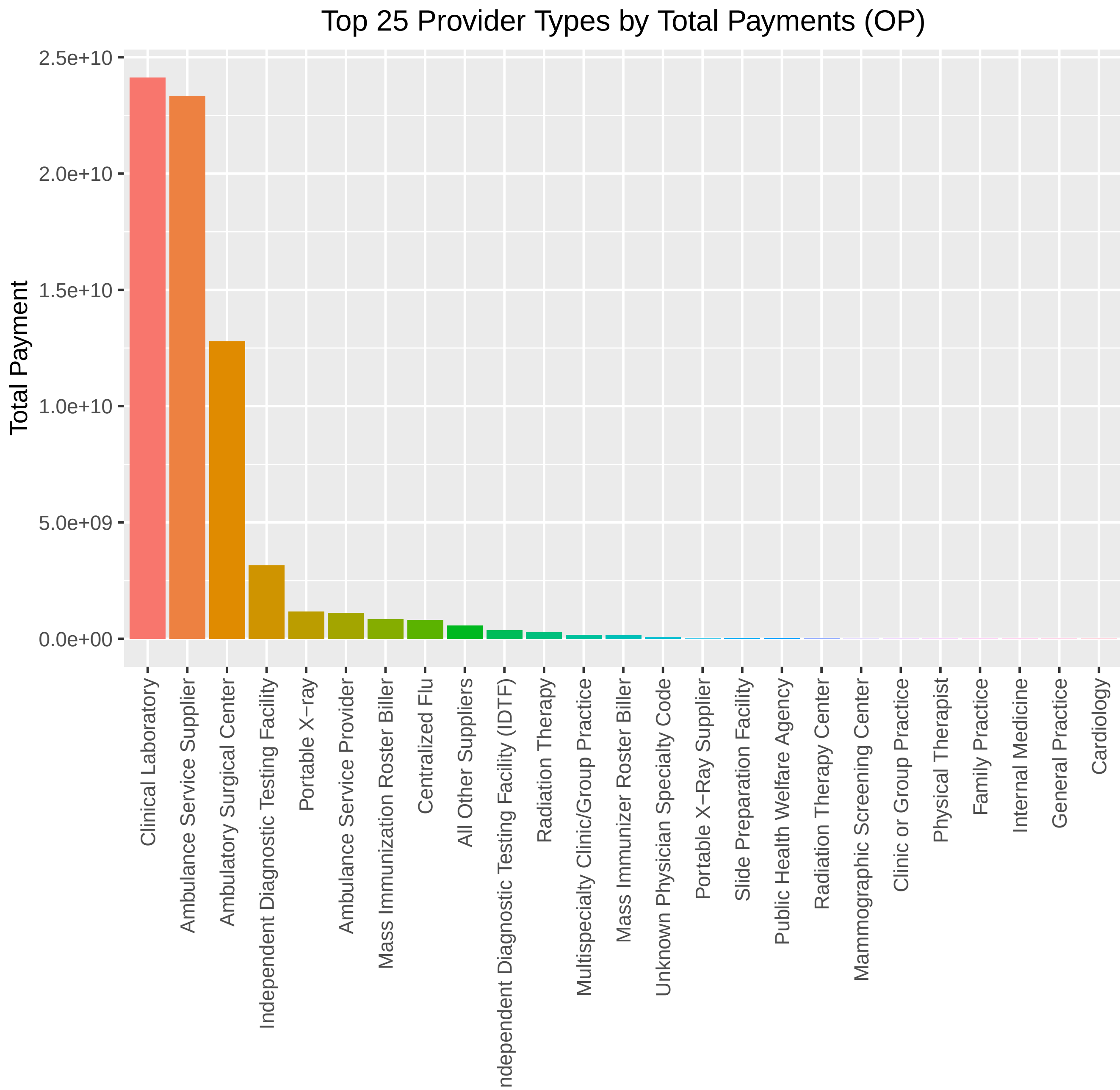
entity.code	mean payment per serv.	mean charge per serv.	mean allowed charge per serv.
I	72.95575	302.9961	95.35338
O	130.93708	668.7736	163.21815

Table 2: Total Payments to Providers

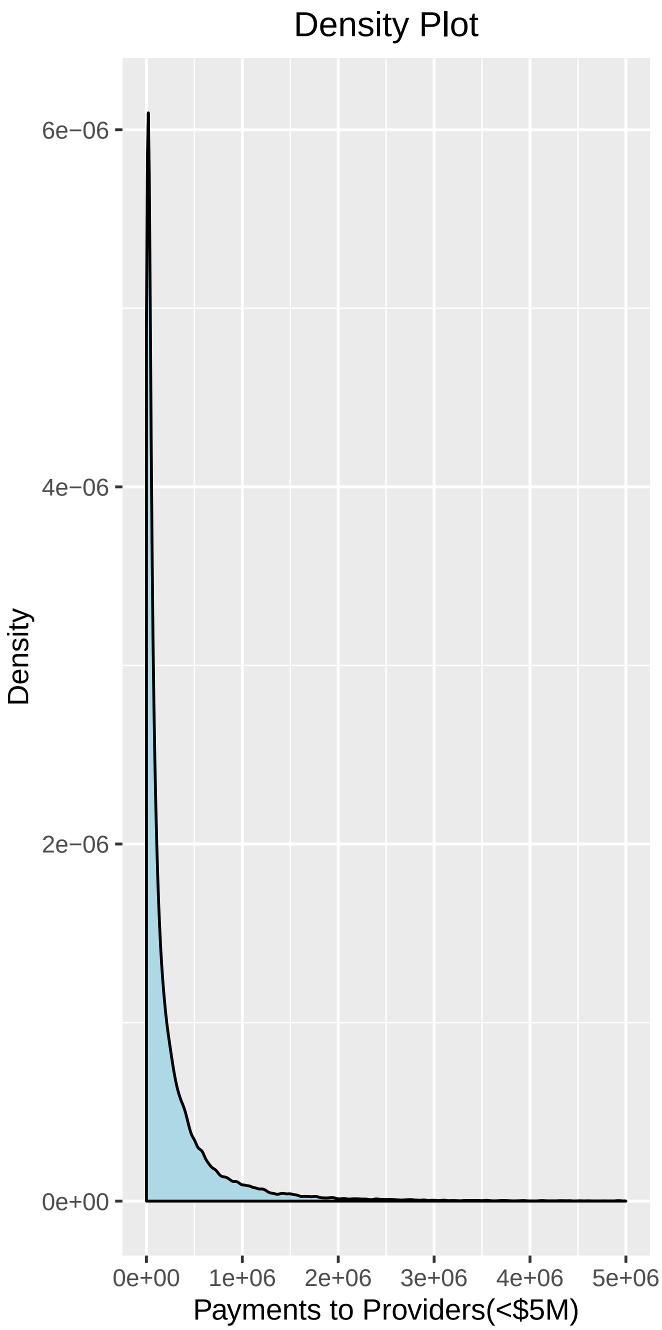
entity.code	median payment per provider	max payment to provider	sum total payment
I	88348.41	52716347	326008816499
O	41326.85	873856208	69115159355

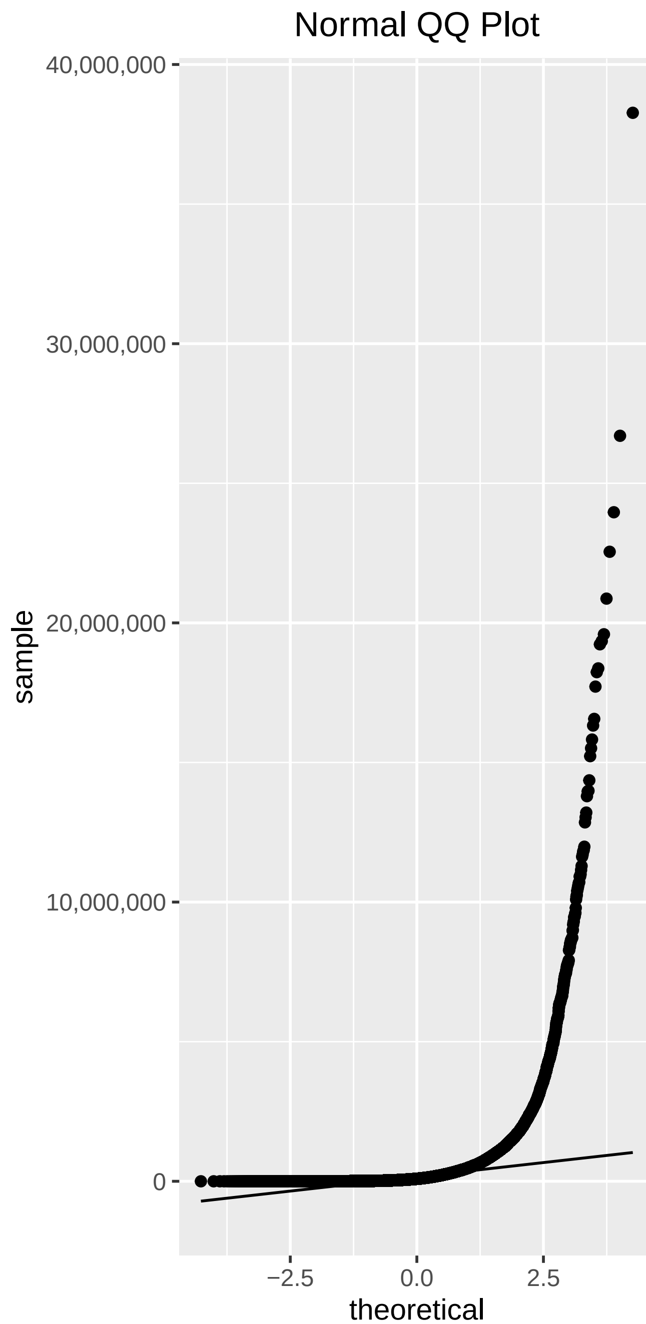
The following graphs present the top 25 provider types ranked by total payments. Note that there is quite a bit of range in total payment, even among the top 25 provider types. (I will investigate the distribution of payments in more detail in the next section.)



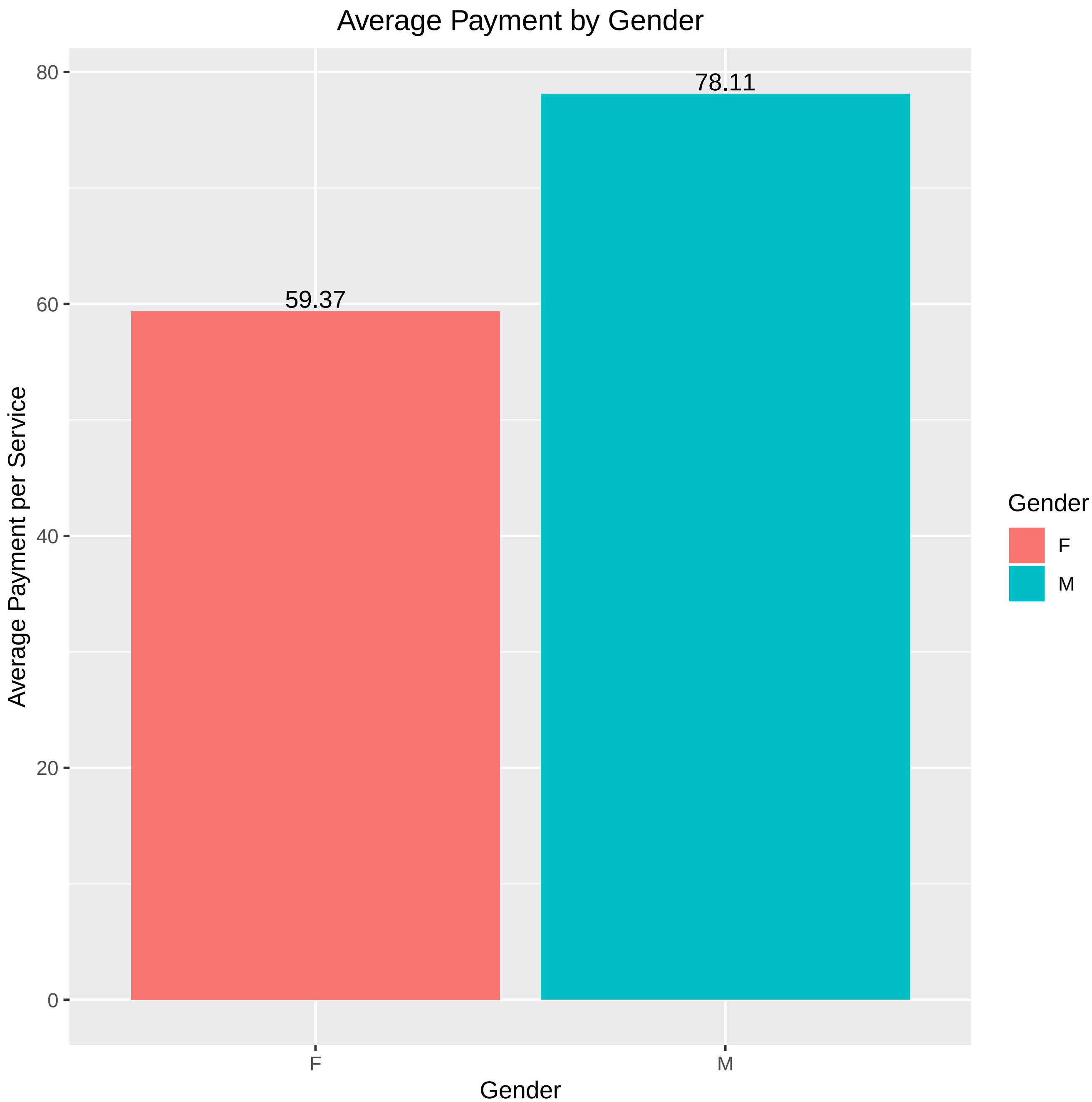


The following graphs illustrate the distribution of total payments made to providers (both IP and OP). The density plot and normal qq plot indicate that the data are heavily skewed to the right. In other words, total payments are concentrated among top providers.





Finally, there seems to be a significant difference in average payment per service for male and female providers:



Based on this limited exploratory analysis, I propose three research questions:

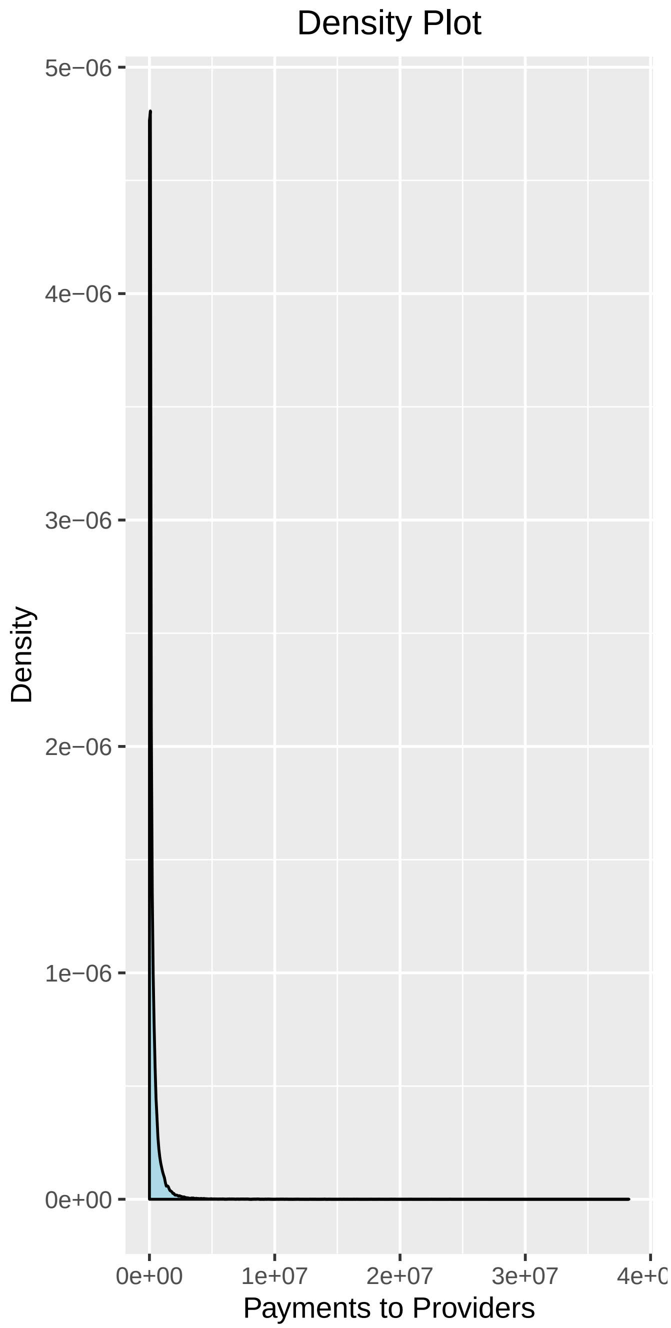
1) How are payments distributed among providers? Is the distribution different for IP and OP?
2) If there is a significant concentration of payments, what might explain that? How does entity type affect payments? How does the concentration of services relate to the concentration of payments?
3) Is there gender inequity in medicare payments? Can that inequality be explained by some other factor, like payment rates for particular healthcare fields (represented by entity type)?

Statistical Analysis

In the following four sections, I will analyze the distribution of payments made to individual and organizational providers, and differences in payment among male and female providers. I will also develop a linear model for predicting payments to providers.

Individual Providers

The density and normal q-q plots for total payments to individual providers show that the data is highly skewed. The density plot is skewed to the right, and the normal qq-plot reveals a distortion at the "top end" of the data. If the data were normal, I would expect to see a bell shaped density plot and a linear normal qq plot.



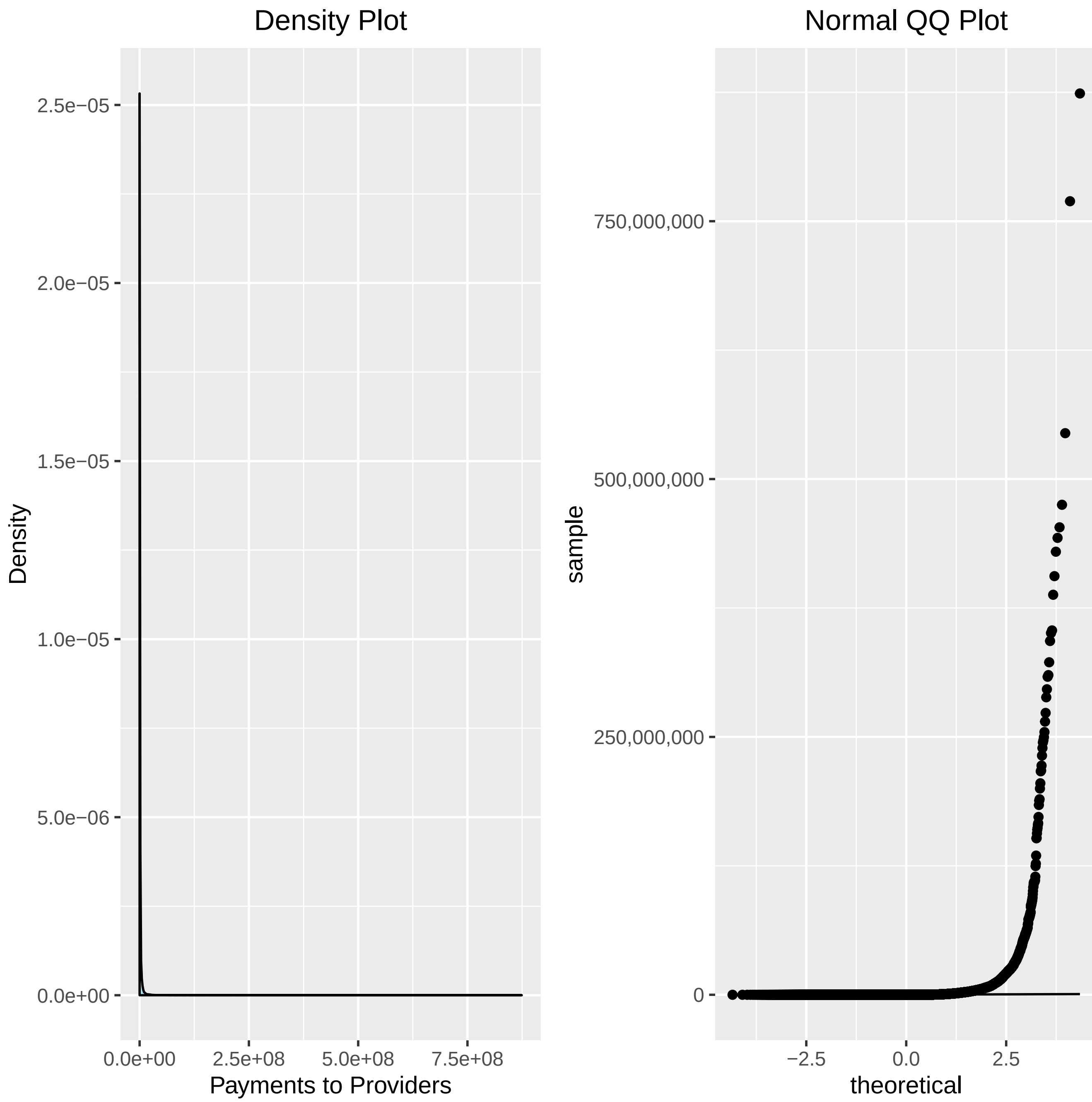
The top individual providers collected the lion's share of payments. The top 10% of providers (ranked by total payments) received 56.5% of payments and the top 50% of providers received 94.9% of payments. In fact, the highest paid provider, who is an ophthalmologist from Florida, collected $52.7 million between 2012 and 2016. The chart below shows total payments and number of services for deciles of individual providers (ranked by total payment). It reveals that the highest paid providers also provided the most services. So, for example, the top 10% of providers received 56.5% of payments and performed 60.68% of services.

Table 3: Payment and Services Distribution by Decile (Individuals)

decile	payment proportion	services proportion	cum. payment proportion	cum. services proportion
90-100	0.5646049	0.6067985	0.5646049	0.6067985
80-90	0.1772965	0.1604471	0.7419013	0.7672457
70-80	0.1038497	0.0917042	0.8457510	0.8589499
60-70	0.0640151	0.0574818	0.9097661	0.9164317
50-60	0.0396023	0.0361476	0.9493684	0.9525794
40-50	0.0242613	0.0223742	0.9736297	0.9749535
30-40	0.0143213	0.0133515	0.9879510	0.9883050
20-30	0.0077446	0.0073027	0.9956956	0.9956077
10-20	0.0034331	0.0033741	0.9991287	0.9989818
0-10	0.0008713	0.0010182	1.0000000	1.0000000

Organizational Providers

The density and normal q-q plots for total payments to organizational providers show that the data is highly skewed. The density plot is skewed to the right, and the normal qq-plot reveals a distortion in the "top end" of the data. If the data were normal, I would expect to see a bell shaped density plot and a linear normal qq plot.



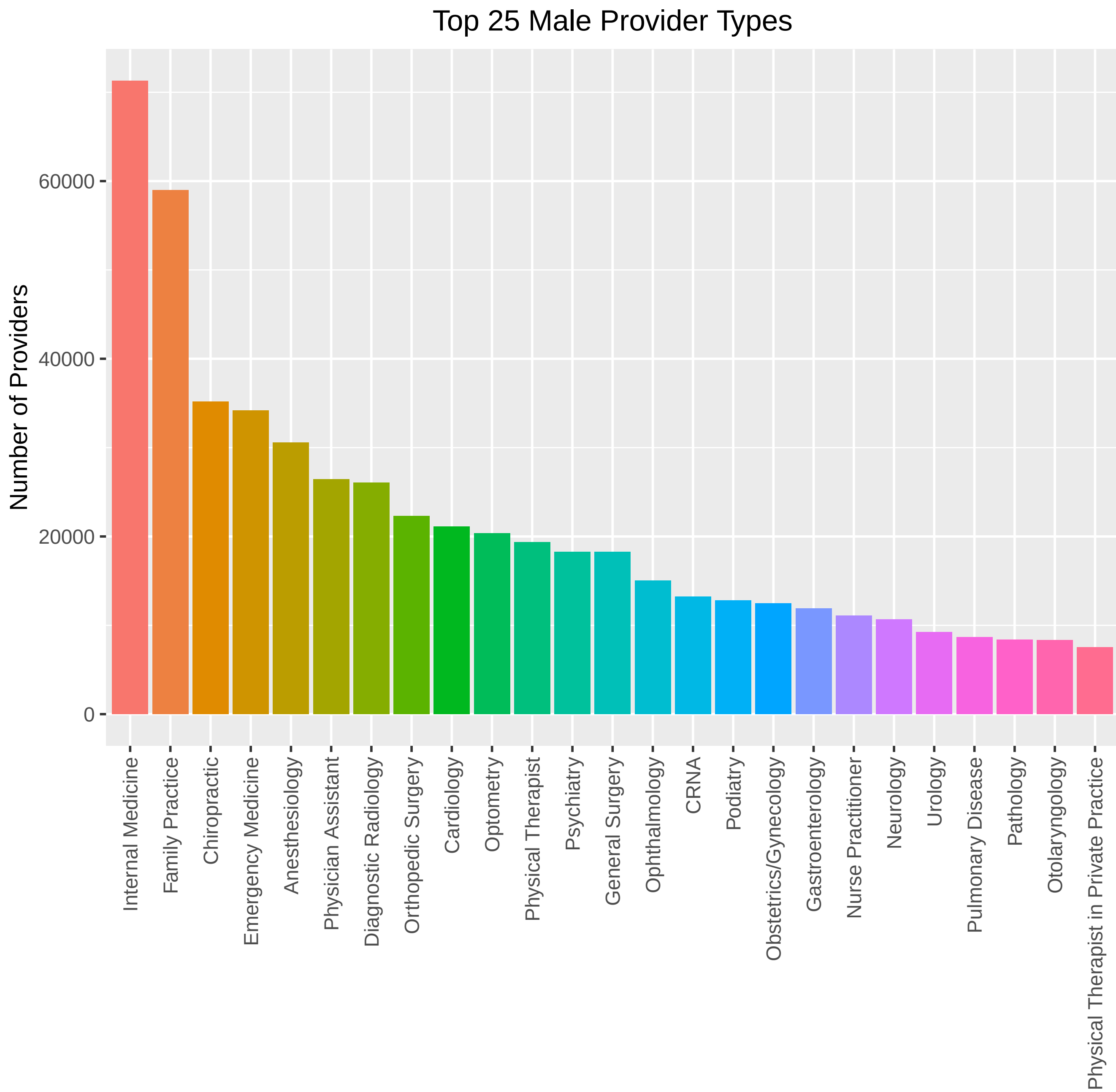
Interestingly, the total payments made to organizational providers are more concentrated at the top than those made to individual providers. The top 10% of providers (ranked by total payments) received 86.7% of payments and the top 50% of providers received 99.2% of payments. The highest paid provider, a clinical laboratory in NC, collected $\$8.7385621 \times 10^8$ between 2012 and 2016. The chart below shows total payments and number of services for deciles of organizational providers (ranked by total payment). It reveals that the highest paid providers also provided the most services. So, for example, the top 10% of providers received 86.7% of payments and performed 88.82% of services. _

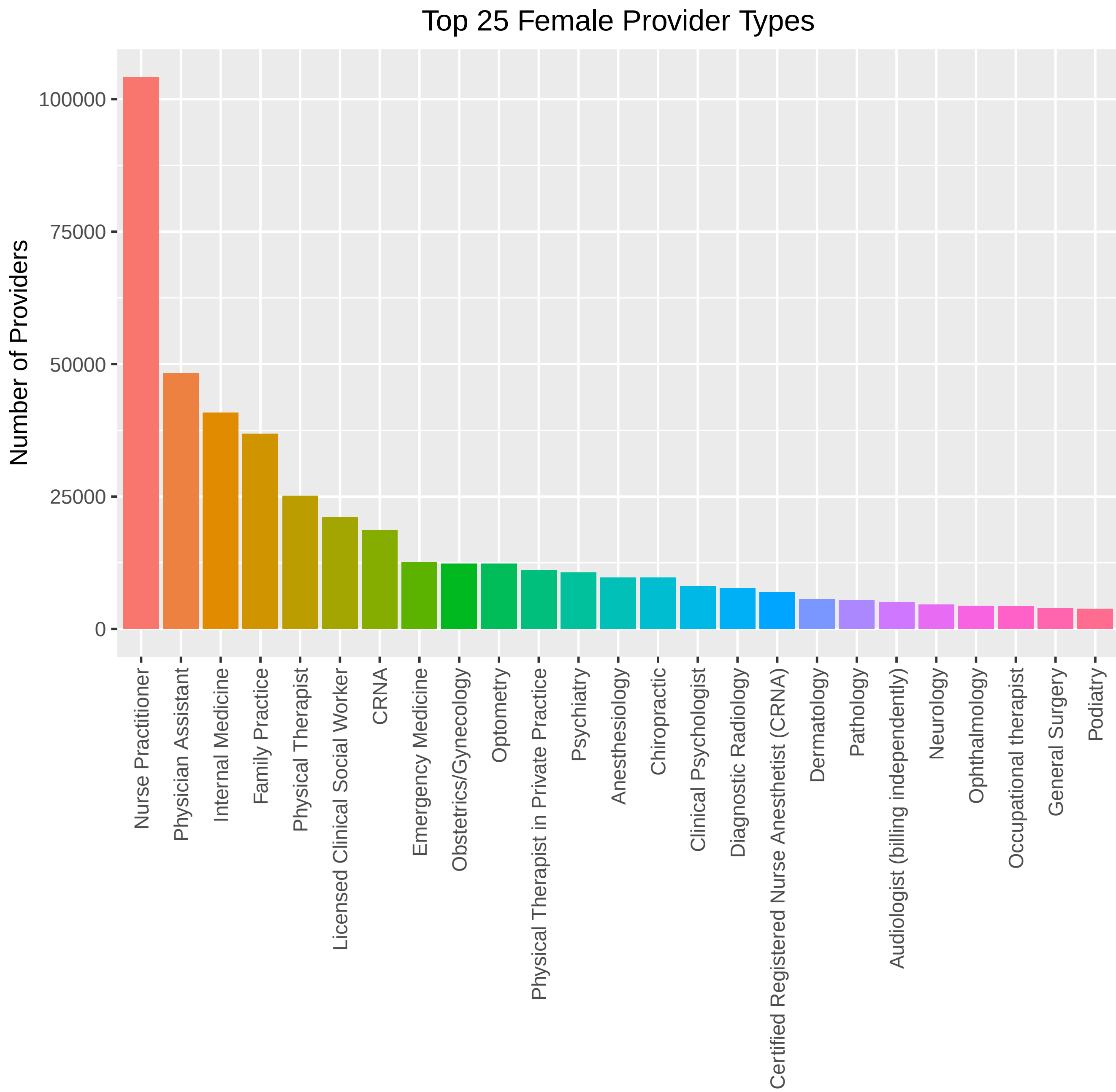
Table 4: Payment and Services Distribution by Decile (Organizations)

decile	payment proportion	services proportion	cum. payment proportion	cum. services proportion
90-100	0.8668794	0.8882215	0.8668794	0.8882215
80-90	0.0892189	0.0681524	0.9560983	0.9563739
70-80	0.0213571	0.0207980	0.9774553	0.9771719
60-70	0.0087807	0.0087911	0.9862360	0.9859630
50-60	0.0053409	0.0052844	0.9915769	0.9912475
40-50	0.0035226	0.0034951	0.9950995	0.9947426
30-40	0.0023629	0.0024345	0.9974624	0.9971771
20-30	0.0014929	0.0015777	0.9989553	0.9987548
10-20	0.0008084	0.0009148	0.9997637	0.9996696
0-10	0.0002363	0.0003304	1.0000000	1.0000000

Gender Disparity

In the exploratory analysis, I noted that the average payment per service is higher for male providers than female providers. In this section, I will consider a common explanation for gender pay inequity: women tend to work in fields that pay less. The following two figures show the top 25 female and male provider types as ranked by the total number of male and female providers in each field:


Top 25 Male Provider Types
Number of Providers
60000
40000
20000
0
Internal Medicine
Family Practice
Chiropractic
Emergency Medicine
Anesthesiology
Physician Assistant
Diagnostic Radiology
Orthopedic Surgery
Cardiology
Optometry
Physical Therapist
Psychiatry
General Surgery
Ophthalmology
CRNA
Podiatry
Obstetrics/Gynecology
Gastroenterology
Nurse Practitioner
Neurology
Urology
Pulmonary Disease
Pathology
Otolaryngology
Physical Therapist in Private Practice

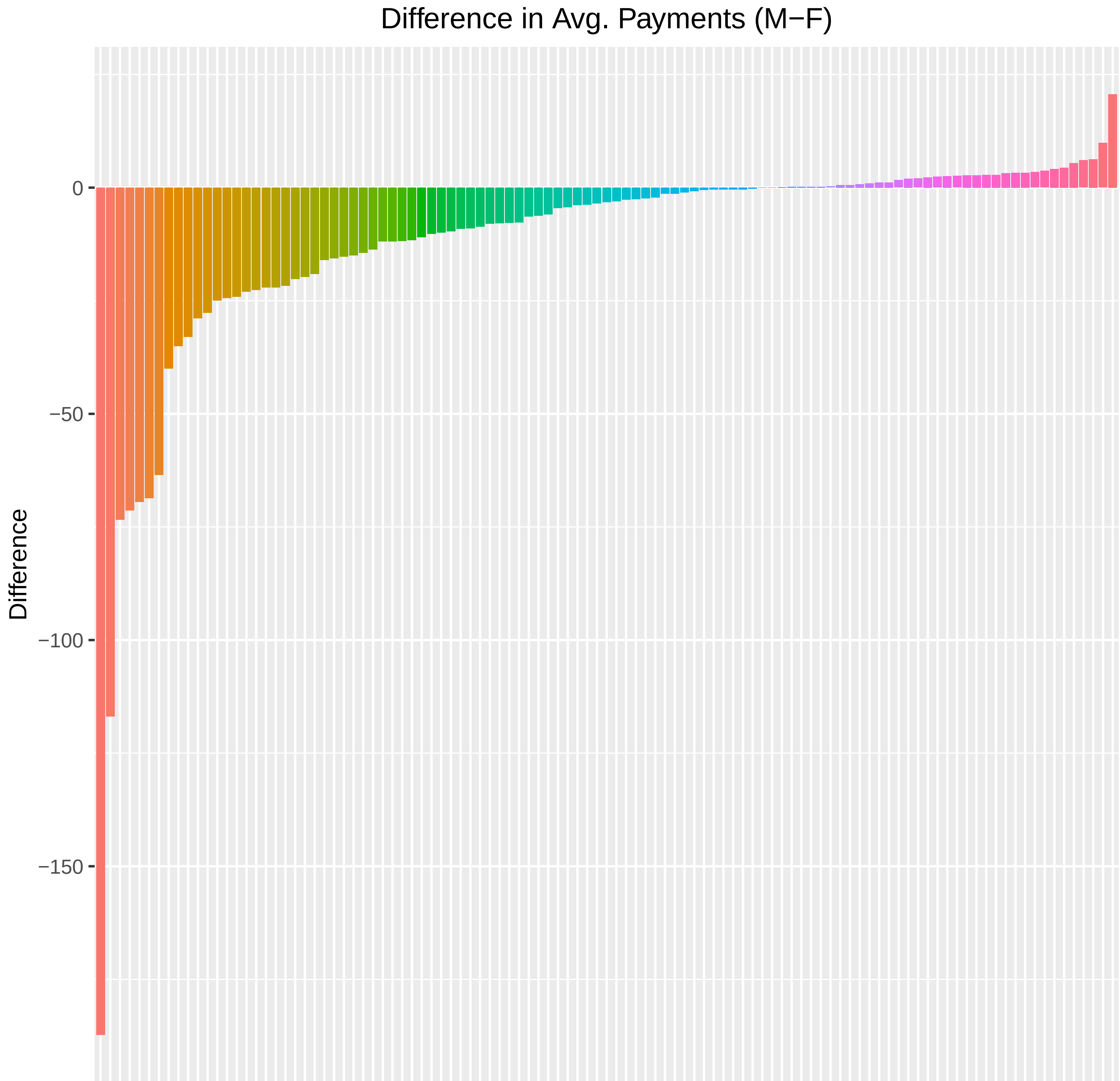


We can quantify the influence of provider type on gender-based pay inequalities by looking at the proportion of total medicare payments made to providers working in specialties that are popular among women and men, respectively.

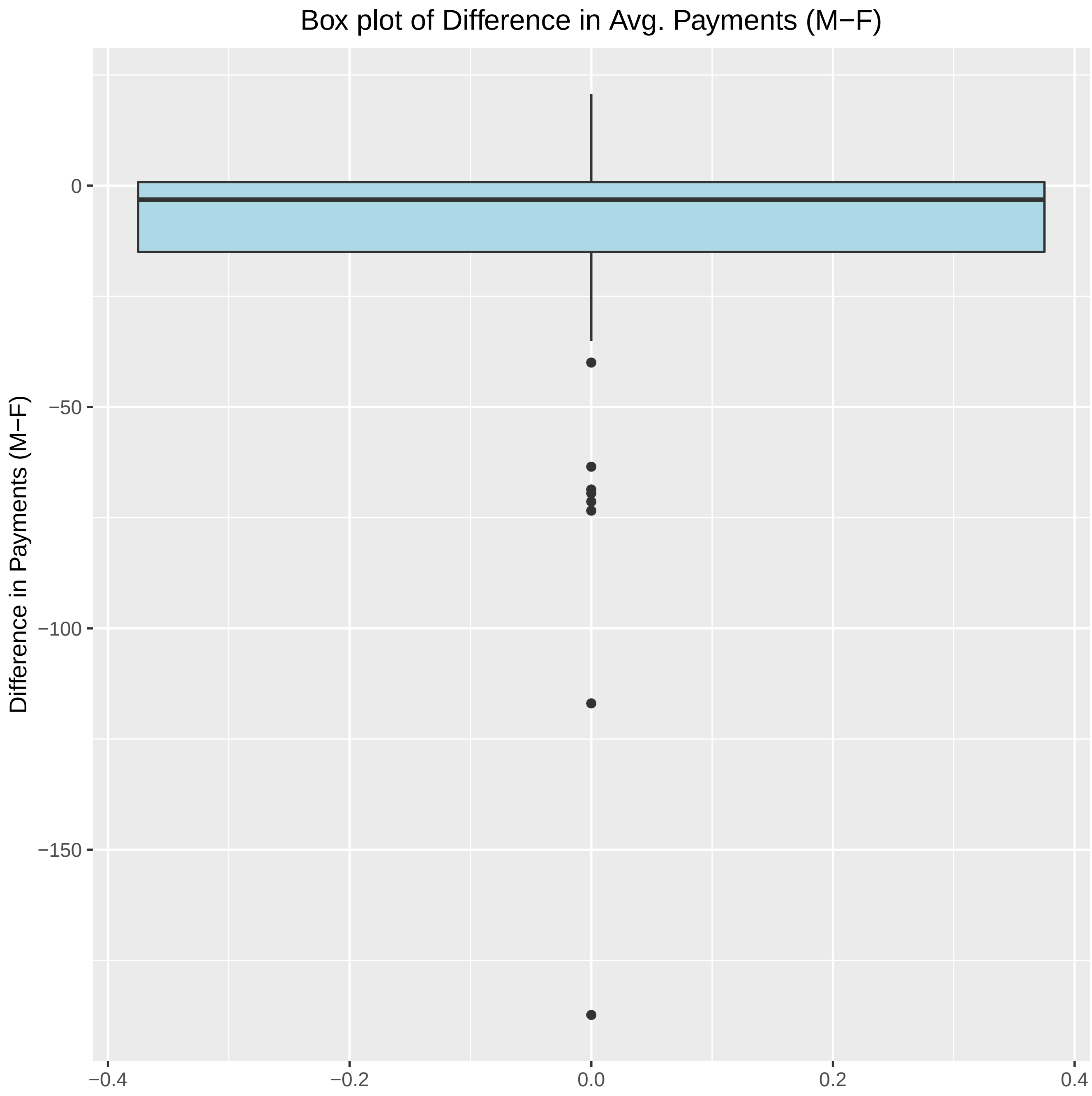
Consider the top 25 female specialties and the top 25 male specialties from the plots above. If we look at the proportion of total payments going to *all* providers in those two "sets" of specialties, we find that the top 25 female specialties account for 61.56% of total payments and the top 25 male specialties account for 74.17% of payments. This supports the claim that gender-based pay inequities are related to the fields in which men and women tend to work.

To give one specific example, the most popular female provider type is Nurse Practitioner (NP): 21.16% of female providers are NPs, while only 1.73% of male providers are NPs. As one might expect, NPs are paid less on average ($43.36 per service) than providers specializing in Internal Medicine (the most popular male provider type), who earn $58.96 per service. Accordingly, the relative prevalence of female NPs may explain lower *average* payment amounts for female providers as a whole.

Yet "provider type" does not entirely explain the pay gap. If we compare "apples to apples" – e.g., average payments to female NPs and average payments to male NPs– a more complicated picture emerges. The following two graphs demonstrate gender pay inequities within the same healthcare field (as represented by provider type). The first graph is a bar plot of the differences between average payments for male and female practitioners who are the same provider type. Negative values indicate that male providers in a certain field receive higher average payments than female providers, while positive values indicate that female providers receive higher average payments than their male counterparts. There are clearly more negative observations than positive observations, which shows that male providers tend to receive higher average payments than their female counterparts.

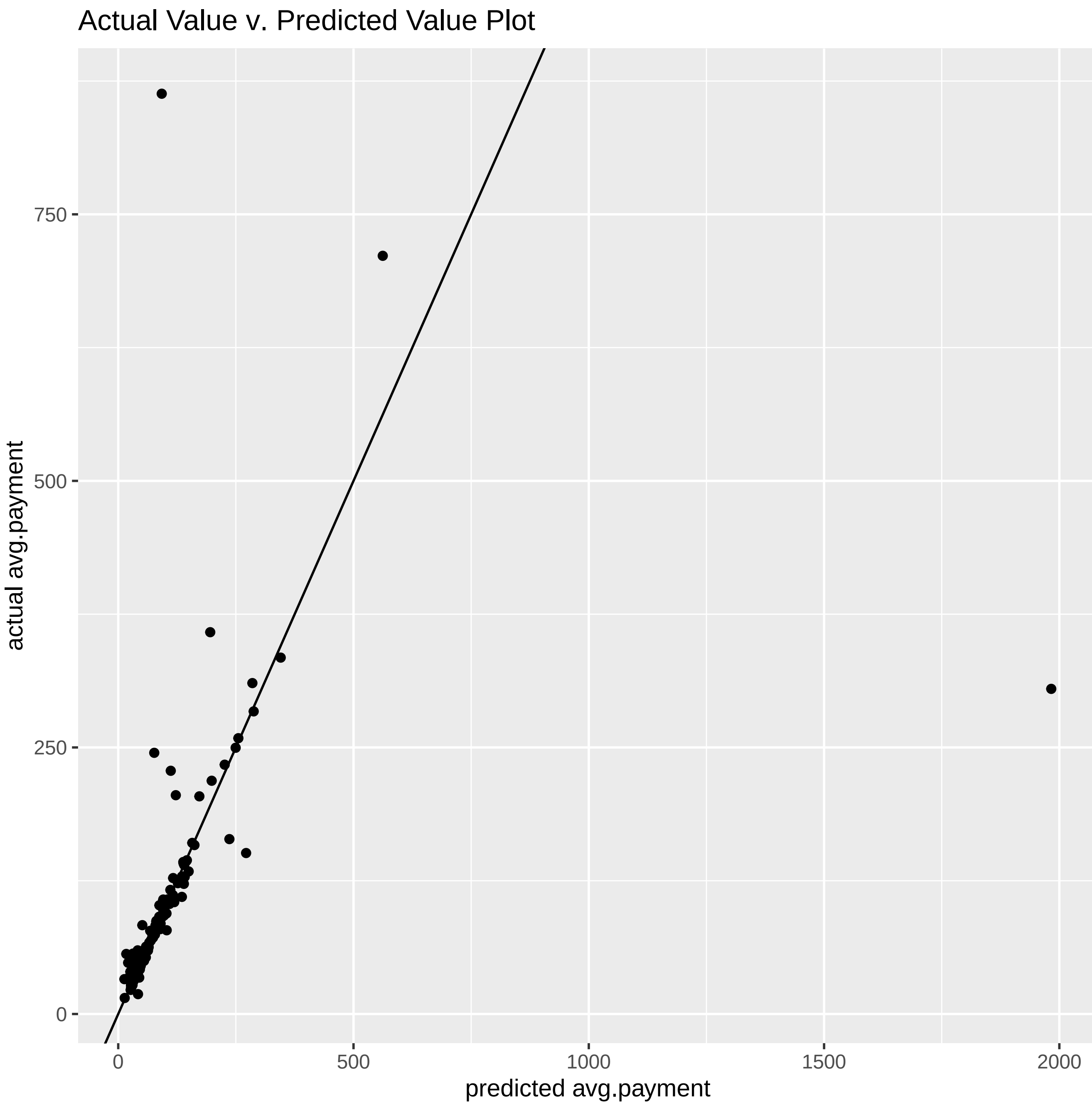


The following graph is a box plot of the differences in average payments to males and females of the same provider type. Women receive $11.9 (95% confidence interval: (-17.02, -6.79) less per service on average than their male colleagues– a 13.14% difference. Female providers tend to receive lower average payments than male providers regardless of provider type. In other words, the data do not support the claim that women are paid less because they tend to work in lower paying fields.



Predictive Model

I used 5% of the "claims" data to build a linear model for "avg.payment" with "entity.code", "prov.type" and "year" as explanatory variables. I then constructed a data frame from the 2017 data consisting of average payments for each provider type and entity code. Finally, I ran the data frame through my model. The following graph shows actual avg. payments for 2017 vs. predicted avg.payments for 2017. The model performs fairly well, though there are a number of outliers. If the model performed perfectly, I would expect all of the observations on the line y=x. Adjusted R^2 is somewhat low: 0.095.



Conclusions and Future Work

This report analyzed the distribution of Medicare payments to individual and organizational providers, the distribution of services, and the distribution of payments with respect to gender. We found that the concentration of payments reflects a similar concentration of services. It seems reasonable for providers to receive more payments for more services. Yet the concentration of payments and services among organizational providers is particularly acute: The top 10% of organizational providers received 86.7% of payments, and the top 50% of providers received 99.2% of payments. This level of concentration raises concerns about monopolies in healthcare industries. It would be worth investigating the pricing practices of major organizational providers further.

The report also noted a significant inequity in medicare payments to male and female providers. Women receive $11.9 (95% confidence interval: (-17.02, -6.79) less per service on average than their male colleagues– a 13.14% difference. Female providers tend to receive lower average payments than male providers regardless of provider type. A difference in average payment amount might also reflect different services rendered by male and female providers of the same type. It would be helpful to include an overview of services (pricing, type, etc.) in an analysis of payment disparities.

Additionally, a linear model for predicting medicare payments was constructed. The actual v. predicted plot revealed that this model was very accurate in some cases, and less accurate in others. Perhaps the model could be improved upon by considering other variables and modelling techniques. For example, it may be helpful to analyze the data for geographic variations in medicare payments and services. Geography could potentially serve as another explanatory variable in the model.