

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Ganguli I, Sheridan B, Gray J, Chernew M, Rosenthal MB, Neprash H. Physician work hours and the gender pay gap in physician pay — evidence from primary care. *N Engl J Med* 2020;383:1349-57.
DOI: 10.1056/NEJMsa2013804

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Ganguli I, Sheridan B, Gray J, et al. Work Effort and the Physician Gender Pay Gap: Evidence from Primary Care.

Table of Contents

Methods	2
Data Source	
Study Sample Definition, Figure S1 . Visit sample criteria	
Covariate and Outcome Definitions, Table S1 . Low Acuity Visit Diagnoses	
Regression Specifications	
Results	9
Table S2 . Gender distribution of primary care specialties, athenahealth data compared to US Physician Compare estimates	
Table S3 . Patient visit characteristics, athenahealth compared to national survey data	
Table S4 . Unadjusted year, day, and visit-level outcomes, by physician gender	
Figure S2 . Adjusted observed time per visit, by physician gender	
Table S5 . Adjusted year, day, and visit-level outcomes including wRVUs, by physician gender	
Table S6 . Adjusted year and day level outcomes using visits with non-missing timestamp data, by physician gender	
Table S7A . Adjusted year-level revenue, visit counts, and time per visit stratified by physician age, by physician gender	
Table S7B . Adjusted visit-level revenue, visit counts, and time per visit stratified by physician age, by physician gender	
Figure S3 . Adjusted observed time per visit by patient and physician gender	
Table S8 . Adjusted visit-level outcomes among patients with 2 or more chronic conditions, by physician gender,	
Table S9 . Adjusted visit-level outcomes among same-day visits, by physician gender	
Table S10 . Adjusted visit-level outcomes by visit type and physician gender	
References	16

Data Source

We used a national sample of de-identified 2011-2017 all-payer claims and electronic health record (EHR) data from athenahealth Inc (“athenahealth”), a cloud-based health care information technology company that provides physician practices with medical billing, practice management, and EHR services. These data include visit-level information derived from submitted claims, including scrambled identifiers (patient, provider, practice), allowed charges, procedure codes, diagnoses codes, dates of service, primary and secondary payer (across all possible payers), and patient demographic characteristics. These claims data can be linked to fields derived from the EHR, including scheduling data (date of appointment scheduling, scheduled visit duration, scheduled visit start time) and timestamps recorded by the EHR during each visit. These data have been used in multiple prior studies.¹⁻⁶

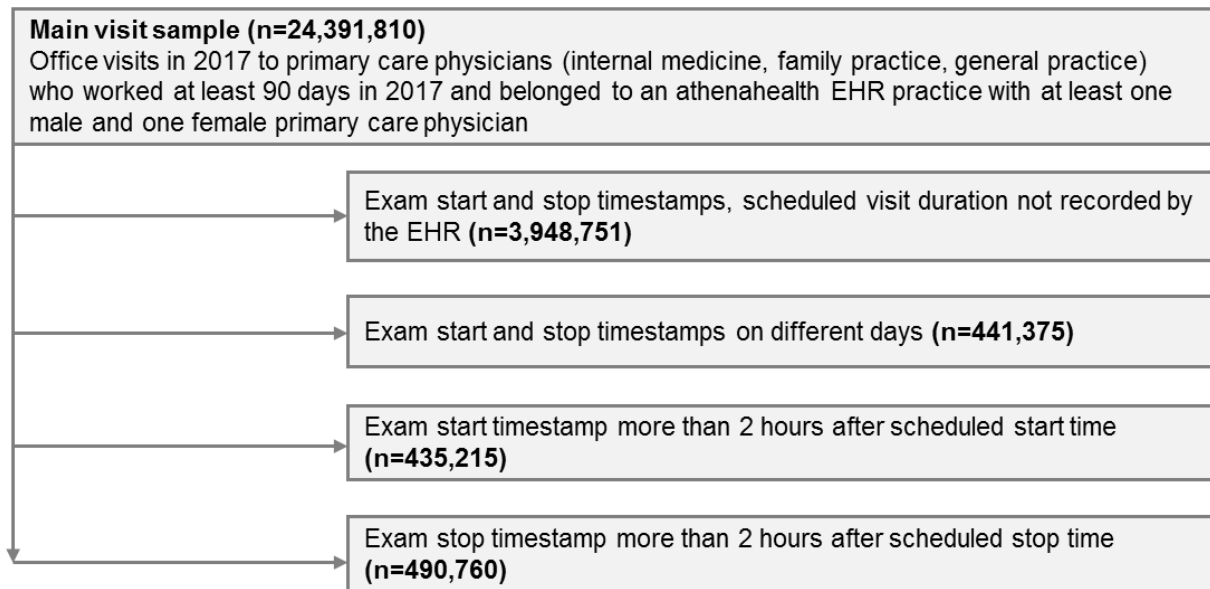
Study Sample Definition

To define our study sample, we employed physician-level, practice-level, and visit-level inclusion criteria. At the physician level, we included any physician (MD or DO; not listed as in training) with a primary care specialty (internal medicine, family practice, general practice) who billed for patient care on at least 90 unique days during 2017 and had age and gender information recorded. At the practice level, we restricted our sample to practices (defined using the highest organization-level identifier available within the athenahealth dataset) with at least one male and one female PCP meeting the criteria described above.

To identify unique visits, we relied on unique combinations of deidentified practice ID, deidentified rendering physician ID, deidentified patient ID, visit date, and scheduled start time (e.g., 9:15am). We used timestamps recorded by the EHR during the visit to determine observed visit length and considered observed visit time to be missing for a given visit if these timestamp data did not meet our quality criteria. Specifically, we considered visit time to be missing for visits with no recorded start or stop time and visits in which the dates for the start and stop timestamps were on different days. We also considered visit length to be missing for visits with a starting timestamp more than two hours different from the scheduled visit time, or those that started within two hours of schedule, but ended more than two hours after their scheduled stop time. See Figure S1 below. We used multiple imputation to impute observed visit length for the 5,316,101 visits that failed one of these quality checks (See more detail under Regression specifications). In a sensitivity analysis, we repeated our main analyses excluding

these visits entirely (Table S6). For secondary analyses of observed time greater than scheduled duration and of percentage of visits scheduled for ≥ 20 minutes, we omitted the 3,948,751 visits without a recorded scheduled visit duration from these analyses specifically.

Figure S1. Visit sample criteria



Covariate and Outcome Definitions

Number of Sessions

To determine scheduled sessions, we defined each physician as having each of ten possible morning (7am-noon) or afternoon (noon-7pm) weekday sessions if they had at least one visit timestamped during that session (e.g., Monday morning) for at least 26 weeks of the study year, then added the number of sessions he or she had in a given weekday (range 0-2) or week (range 1-10).

Chronic Condition Count

We created 27 chronic condition indicators based on patients' history of diagnoses (ICD-10 codes) and procedures (CPT/HCPCS codes) similar to the algorithm used by the Chronic Conditions Data Warehouse.⁷ When possible, we used the reference period (ranging from six months to three years) specified in the algorithm. From the 27 condition-specific measures, we constructed a summary measure counting the number of conditions present for any given

patient at any given visit. Athenahealth data likely does not include complete claims for all patients, so our measures may under-capture presence of chronic conditions within our sample.

Visit for Low Acuity Condition

We counted a visit as addressing a low acuity condition based on the primary ICD-10 diagnosis code linked to the visit (See Table S1).

Table S1. Low Acuity Visit Diagnoses

Low-Acuity Condition	ICD-10
Upper respiratory infection (including acute nasopharyngitis, laryngitis, tracheitis)	J00.xx, J04.xx, J06.xx
Sinusitis	J01.xx
Bronchitis	J20
Pharyngitis	J02.xx
Otitis media	H65.xx
Otitis externa	H60.xx
Conjunctivitis	H10.xx
Urinary tract infection	N39.0
Allergic rhinitis	J30.xx
Influenza	J09, J10, J11
Unspecified viral infection	B34.xx

Based on Ganguli I et al. Ann Int Med. 2020, originally adapted from Reid RO et al. J Gen Intern Med. 2013.^{8,9}

Visit Revenue

To quantify visit revenue, we summed allowed charges for all services rendered on the day of visit at the year, day, and visit levels. This dollar figure reflected the negotiated amount paid by each payer for each service.

Diagnosis Count

At the visit level, we counted the total number of diagnoses (unique ICD-10 codes) recorded on all claims submitted for each visit.

Order Count

The athenahealth EHR tracks orders placed for lab studies, imaging tests, prescription drugs, and referrals to other clinicians (e.g., physical therapy, specialist referrals). In our visit level analysis, we counted the total number of orders placed for a given patient by the rendering clinician on the day of the visit.

Observed Visit Time

The athenahealth EHR structures patient visits in six stages: patient check-in, intake, exam, check-out, sign-off, and post-visit documentation. To measure observed visit time, we used timestamps recorded during the exam stage, quantifying observed exam length as the number of integer minutes from the first exam-stage timestamp to the last exam-stage timestamp.

Clinical context: Timestamps capture a given electronic health record user's actions in an EHR and are specifically linked to that user via their log-in. The athenahealth EHR timestamps examined for this study capture clinicians and staff advancing the patient encounter through sequential stages both to access and complete tasks relevant to that stage and to communicate the patient's status to other team members by updating this status in the schedule view of the EHR.

Hypothetical patient encounter: Once a patient arrives, a front desk staff member completes the check-in (e.g., confirming insurance), then a medical assistant takes the patient to an exam room for intake (e.g., vital signs, medication reconciliation). Once intake is complete (schedule view status: "Ready for provider"), the physician can then click "Go to Exam" to start the visit (i.e., exam start). This exam stage allows the physician to run through visit-specific tasks (e.g., draft history of present illness, document exam, place and sign orders, draft patient instructions for the after-visit summary required by meaningful use regulations, and write a follow-up plan). At the end of the visit, the physician closes the exam stage (i.e., exam stop) to advance the encounter to the checkout stage. This action updates the patient's status to "Ready for checkout" in the schedule view, signaling to the checkout staff that the patient is on their way. This action also allows the checkout staff to start the checkout phase in order to access checkout materials (e.g., print the after visit summary or schedule a follow-up visit suggested by the physician).

Physician pre- and post-visit work: Before the visit, the physician can use the review dashboard to review the patient's record, including even the vitals and reason for visit being collected during intake. This occurs outside of the patient encounter – that is, without opening the exam stage. After closing the exam stage, the physician can keep the EHR open and complete encounter tasks such as finalizing documentation and billing, or close the EHR and return to the encounter later. This work is logged as "post-visit documentation."

Measure validation and refinement: As described in the study sample validation, we considered this measure to be missing if a visit's timestamp data did not meet our quality criteria. We also took the following analytic steps to clean the data and to confirm our assumptions:

Though the physician does not need to enter the exam stage to review a patient's chart, it is theoretically possible for her to enter this stage any time after the patient checks in, which might cause overestimation of visit length. In our data, this would manifest as the start of the exam stage occurring before the end of the intake stage. We found that this was an equally rare occurrence for male and female PCPs (3.1% of visits by male PCPs vs 2.9% of visits by female PCPs, difference -0.001% [95%CI -0.004% – 0.001%]).

If the physician fails to close the exam stage at the end of the visit and the patient still goes to check out, the checkout staff might start checkout then close the exam themselves, manifesting as checkout starting before the exam is over. This was also equally rare (0.05% of visits) for both male and female PCPs (Difference 0.001%, 95%CI [-0.02- 0.02]).

To further test for gender differences in timestamp data quality that might bias our results, we compared the percentage of visits with anomalous timestamps (see Figure S1) by physician gender, including practice fixed effects. Male and female PCPs had similar rates of these visits: 6.4% vs 6.5%, (Difference -0.1%, 95%CI [-0.4 - 0.2]).

During a visit, patients may interact with multiple clinicians (i.e., a medical assistant and a physician), generating multiple timestamps (i.e., rows) in the athenahealth EHR metadata. In these instances, we used only the record in which the deidentified EHR user ID matched the ID of the rendering physician on the claim.

To account for the common practice of double-booking visits in primary care, we identified and allocated overlapping visit-minutes evenly between affected visits. For each minute in a physician's day, we observed the number of exams that appeared to occur in that minute. We apportioned each overlapping minute equally between all affected exams. For example, if two exams in a physician's day both included the minute from 9:00 to 9:01, each exam received 30 seconds toward the total observed duration.

Visit Time Greater than Scheduled

This variable was constructed by subtracting scheduled visit duration from observed visit time. Visits for which scheduled visit duration was not recorded in the data (16.2% of visits, see Figure S1) were omitted for analyses of this outcome specifically.

Work Relative Value Units (wRVUs)

In a sensitivity analysis, we summed the number of wRVUs associated with each service rendered at the visit, day, and yearly levels. wRVUs varied by service type (i.e., Current Procedural Terminology [CPT] code), but not by payer type.

Regression Specifications

Main analyses - We estimated ordinary least squares regressions at the annual, day, and visit levels. At the annual level, our main analysis estimated the following specification:

$$Y_{kp} = \beta_1 FemalePCP_k + PhysChar_k + \gamma_p + \epsilon_{kp}$$

indexing physician k at practice p . Y_{kp} were the primary and secondary outcomes described in the main text. $FemalePCP_k$ was an indicator for whether the PCP was female. $PhysChar_k$ was a vector of physician characteristics including physician age (continuous), degree (MD or DO), specialty (internal medicine, family practice, general practice), and session count. γ_p represented practice fixed effects.

At the day level, our main analysis estimated the following specification:

$$Y_{kpt} = \beta_1 FemalePCP_k + PhysChar_k + Weekday_t + \gamma_p + \epsilon_{kpt}$$

indexing physician k at practice p on day t . Y_{kpt} were the primary and secondary outcomes described in the main text. $FemalePCP_k$ was an indicator for whether the PCP was female. $PhysChar_k$ was a vector of physician characteristics including physician age (continuous), degree (MD or DO), specialty (internal medicine, family practice, general practice), and session count on the given day. $Weekday_t$ represented day-of-the-week fixed effects and γ_p represented practice fixed effects.

At the visit level, our main analysis estimated the following specification:

$$Y_{ijkp} = \beta_1 FemalePCP_k + VisitChar_i + PatChar_j + PhysChar_k + \gamma_p + \epsilon_{ijkp}$$

indexing visit i for patient j to physician k at practice p . Y_{ijkp} were the primary and secondary outcomes described in the main text. $FemalePCP_k$ was an indicator for whether the PCP was female. $VisitChar_i$ included patient age category, chronic condition count, primary payer (Medicare fee-for-service, Medicare Advantage, Medicaid, dual-eligible, commercial insurance, uninsured, other), day-of-week dummies, indicator for same-day visit, and an indicator for

whether the patient had seen that physician previously (since 2011). $PatChar_j$ included patient gender, race/ethnicity, and marital status. $PhysChar_k$ included physician age, degree (MD or DO), and specialty (internal medicine, family practice, general practice). γ_p represented practice fixed effects.

Multiple imputation of missing data: For visits in which time stamp data were not recorded (16.2% of our main visit sample) or did not meet our quality standards (5.6%, see Figure S1), we relied on model-based multiple imputation to impute observed visit time.¹⁰ Specifically, we used the remaining 19,075,709 visits with non-anomalous timestamps to estimate visit time as a function of visit, patient, and physician characteristics with practice fixed effects:

$$ObservedVisitTime_{ijkp} = VisitChar_i + PatChar_j + PhysChar_k + \gamma_p + \epsilon_{ijkp}$$

indexing visit i for patient j to physician k at practice p . $VisitChar_i$ included patient age category, chronic condition count, primary payer, day-of-week dummies, evaluation and management intensity level, count of diagnoses recorded, an indicator for low-acuity visit, and an indicator for whether the patient had seen that physician previously (since 2011). $PatChar_j$ included patient gender, race/ethnicity, and marital status. $PhysChar_k$ included physician age, gender, degree, and specialty. γ_p represented practice fixed effects.

Results

Table S2. Gender distribution of primary care specialties, athenahealth data compared to US Physician Compare estimates

Physician Characteristic	athenahealth (2017)	Physician Compare (2017)
<i>Gender, %</i>		
Female	36.4	34.1
Male	63.7	65.9
<i>Credential, %</i>		
MD	85.8	90.2
DO	14.3	9.8
<i>Specialty, %</i>		
Internal Medicine	53.2	52.1
Family Practice	46.1	45.1
General Practice	0.8	2.9
<i>Census Region, %</i>		
Northeast	18.5	20.1
Midwest	24.1	24.4
South	47.6	33.9
West	9.7	21.6

Source: Authors' analysis of athenahealth data and publicly available Physician Compare database.¹¹ Note: Analysis of Physician Compare database limited to physicians specializing in internal medicine, family practice, and general practice. Percentages may not sum to one due to rounding.

Table S3. Patient visit characteristics, athenahealth compared to national survey data

Patient Visit Characteristic	athenahealth (2017)	NAMCS (2016)
<i>Payer, %</i>		
Medicare	39.0	27.5
Medicaid	7.8	10.7
Dual	4.7	2
Commercial	45.1	50.9
No Insurance	2.0	2
Other Payer	1.4	9
<i>Age Category, %</i>		
<15	3.2	5.3
15-24	4.5	6.2
25-44	15.9	21.4
45-64	34.3	37.3
65+	42.1	29.8
<i>Sex, %</i>		
Female	56.3	55.9
<i>Race/Ethnicity, %</i>		
White, non-Hispanic	75.1	69.3
Non-White	11.8	25.5
Other/Unknown	13.1	5.2
<i>Chronic Conditions, #</i>		
0	49.1	29.3
1	22.2	22
2+	28.8	48.7
<i>Geographic Region, %</i>		
Northeast	18.0	17.6
Midwest	21.5	28.8
South	52.4	38.3
West	8.2	15.4

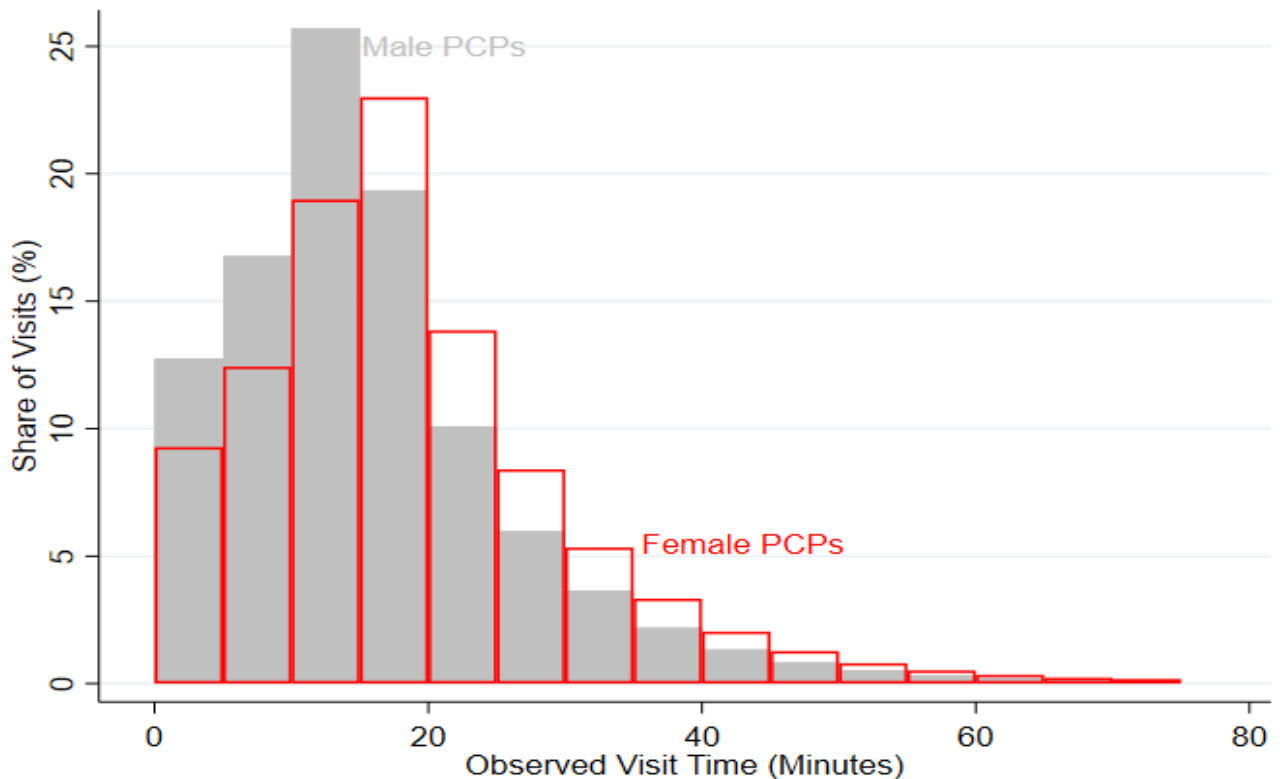
Source: Authors' analysis of visit-level patient characteristics from athenahealth data and publicly available National Ambulatory Medical Care Survey data (NAMCS). Note: Percentages may not sum to one due to rounding.

Table S4. Unadjusted year, day, and visit-level outcomes, by physician gender

	Male PCPs (N=5,284)	Female PCPs (N=3,018)	Difference	95% CI for Difference	Difference, %
A. Year-level					
Allowed charges, \$	360,820.8	316,101.9	-44,718.9	-60,525.7 - -28,912.1	-12.4
Visits, no.	3,077.9	2,693.2	-384.8	-479.3 - -290.3	-12.5
Days in clinic	204.5	195.7	-8.8	-11.5 - -6.1	-4.3
Observed visit time, minutes	46,908.2	47,562.2	654.0	-844.8 - 2,152.9	1.4
B. Day-level					
	(N=1,085,623)	(N=585,808)			
Allowed charges, \$	1,794.5	1,607.6	-186.9	-250.8 - -123.0	-10.4
Visits, no.	15.2	13.5	-1.6	-1.9 - -1.3	-10.5
Observed visit time, min.	231.3	239.5	8.2	3.9 - 12.6	3.5
C. Visit-level					
	(N=16,422,457)	(N=7,969,353)			
Allowed charges, \$	117.0	117.8	0.8	-2.8 - 4.4	0.7%
Diagnoses documented, no.	3.5	3.6	0.1	0.1 - 0.2	2.9%
Orders placed, no.	2.6	3.1	0.5	0.4 - 0.6	19.2%
Observed visit time, min.	15.3	17.7	2.4	2.2 - 2.7	15.7%
Visit time > scheduled, min. ^a	-1.0	0.9	1.9	1.7 - 2.1	190.0%

^aCalculated by subtracting scheduled visit duration from observed visit time, among visits for which scheduled duration available.

Figure S2. Adjusted observed time per visit, by physician gender



This histogram shows the visit-level distribution of observed visit time among male and female PCPs.

Table S5. Adjusted year, day, and visit-level outcomes including wRVUs, by physician gender

	Male PCPs (N=5,284)	Female PCPs (N=3,018)	Difference	95% CI for Difference	Difference, %
A. Year-level					
Allowed charges, \$	358,795.1	319,652.0	-39,143.2	-53,523.0 - -24,763.4	-10.9%
wRVU	3705.4	3352.7	-352.7	-445.4 - -260.1	-9.5%
Visits, no.	3,058.2	2,727.7	-330.5	-406.6 - -254.3	-10.8%
Days in clinic	203.3	197.9	-5.3	-7.7 - -3.0	-2.6%
Observed visit time, min.	46,709.2	47,910.6	1,201.3	184.7 - 2,218.0	2.6%
B. Day-level					
	(N=1,085,623)	(N=585,808)			
Allowed charges, \$	1,792.3	1,611.6	-180.7	-246.8 - -114.7	-10.1%
wRVU	18.5	16.9	-1.7	-2.2 - -1.4	-9.2%
Visits, no.	15.2	13.6	-1.6	-1.9 - -1.3	-10.5%
Observed visit time, min.	231.5	239.2	7.7	3.6 - 11.8	3.3%
C. Visit-level					
	(N=16,422,457)	(N=7,969,353)			
Allowed charges, \$	117.4	116.9	-0.5	-4.3 - 3.2	-0.4%
wRVU	1.2	1.2	0	-0.0 - 0.0	0.0%
Diagnoses documented, no.	3.4	3.7	0.2	0.2 - 0.3	5.9%
Orders placed, no.	2.6	3.1	0.5	0.4 - 0.5	19.2%
Observed visit time, min.	15.3	17.6	2.4	2.1 - 2.6	15.7%
Visit time > scheduled, min. ^a	-0.9	0.8	1.8	1.5 - 2.0	188.9%

At the year and day levels, we adjusted for physician age, degree, specialty, and sessions worked per week. At the visit level, we additionally adjusted for patient characteristics (age, gender, race/ethnicity, marital status, number of chronic conditions, primary insurer, whether the patient was new to the physician), and visit characteristics (visit type, diagnosis type, advance vs same-day). ^aCalculated by subtracting scheduled visit duration from observed visit time, among visits for which scheduled duration available.

Table S6. Adjusted year and day level outcomes using visits with non-missing timestamp data, by physician gender

	Male PCPs	Female PCPs	Difference	95% CI for Difference	Difference, %
A. Year-level					
	(N=5,284)	(N=3,018)			
Allowed charges, \$	285,372.0	259,358.8	-26,013.2	-34,477.0 - -17,549.5	-9.1%
Visits, no.	2,385.4	2,144.3	-241.1	-301.0 - -181.2	-10.1%
Days in clinic, no.	174.9	173.7	-1.2	-3.2 - 0.7	-0.7%
Observed visit time, min.	38,397.3	39,599.8	1,202.5	337.8 - 2,067.2	3.1%
B. Day-level					
	(N=1,085,623)	(N=585,808)			
Allowed charges, \$	1,416.4	1,291.5	-124.9	-159.3 - -90.5	-8.8%
Visits, no.	11.8	10.6	-1.2	-1.4 - -1.0	-10.2%
Observed visit time, min.	190.4	197.5	7.1	3.5 - 10.7	3.7%
C. Visit-level					
	(N=12,653,467)	(N=6,422,233)			
Allowed charges, \$	119.9	120.4	0.5	-1.3 - 2.4	0.10%
Diagnoses documented, no.	3.8	4.0	0.2	0.2 - 0.3	5.10%
Orders placed, no.	3.1	3.6	0.5	0.4 - 0.6	15.60%
Observed visit time, min.	16.1	18.5	2.4	2.1 - 2.7	15.30%
Visit time > scheduled, min. ^a	-0.9	0.8	1.8	1.5 - 2.0	237.50%

In this a sensitivity analysis, we repeated our main analyses using only the 19,075,709 visits meeting the timestamp quality criteria shown in Figure S1. ^aCalculated by subtracting scheduled visit duration from observed visit time (no missing data).

Table S7A. Adjusted year-level revenue, visit counts, and time per visit stratified by physician age, by physician gender

Physician age, years	Male PCPs	Female PCPs	Difference	95% CI for Difference	Difference, %
25-44	(N=1,352)	(N=1,368)			
Allowed charges, \$	331,473.0	303,328.1	-28,144.9	-63,201.8 - 6,912.0	-8.5%
Visits, no.	2,726.8	2,493.4	-233.4	-345.5 - -121.3	-8.6%
Days in clinic	194.6	190.3	-4.3	-8.3 - -0.3	-2.2%
Observed visit time, min.	42,453.4	44,387.1	1,933.7	223.6 - 3,643.7	4.6%
45-64	(N=2,984)	(N=1,503)			
Allowed charges, \$	385,053.2	336,601.2	-48,452.0	-68,496.2 - -28,407.7	-12.6%
Visits, no.	3,319.7	2,902.5	-417.2	-526.6 - -307.8	-12.6%
Days in clinic	209.8	202.9	-7.0	-10.1 - -3.8	-3.3%
Observed visit time, min.	50,272.5	50,887.1	614.7	-935.8 - 2,165.1	1.2%
≥65	(N=940)	(N=134)			
Allowed charges, \$	320,528.1	259,401.5	-61,126.6	-113,035.5 - -9,217.8	-19.1%
Visits, no.	2,804.9	2,451.7	-353.2	-852.2 - 145.9	-12.6%
Days in clinic	199.3	189.4	-9.9	-19.0 - -0.7	-5.0%
Observed visit time, min.	42,745.1	41,029.1	-1,716.0	-6,227.8 - 2,795.9	-4.0%

Table S7B. Adjusted visit-level revenue, visit counts, and time per visit stratified by physician age, by physician gender

Physician age, years	Visits to Male PCPs	Visits to Female PCPs	Difference	95% CI for Difference	Difference, %
25-44	(N=3,728,552)	(N=3,369,064)			
Allowed charges, \$	122.5	120.4	-2.0	-11.9 - 7.8	-1.6%
Diagnoses documented, no.	3.3	3.5	0.2	0.1 - 0.3	6.1%
Orders placed, no.	2.6	3.1	0.5	0.4 - 0.6	19.2%
Observed visit time, min.	15.5	17.9	2.4	1.8 - 2.9	15.5%
Visit time > scheduled ^a	-0.9	1.0	1.9	1.3 - 2.4	211.1%
45-64	(N=10,025,949)	(N=4,245,381)			
Allowed charges, \$	115.7	116.6	0.9	-3.5 - 5.2	0.8%
Diagnoses documented, no.	3.5	3.7	0.2	0.2 - 0.3	5.7%
Orders placed, no.	2.6	3.1	0.5	0.4 - 0.6	19.2%
Observed visit time, min.	15.2	17.6	2.4	2.1 - 2.7	15.8%
Visit time > scheduled ^a	-0.9	0.7	1.7	1.3 - 2.0	188.9%
≥65	(N=2,644,487)	(N=320,625)			
Allowed charges, \$	113.4	111.9	-1.6	-16.6 - 13.5	-1.4%
Diagnoses documented, no.	3.5	3.8	0.3	0.1 - 0.5	8.6%
Orders placed, no.	2.5	2.9	0.4	0.1 - 0.7	16.0%
Observed visit time, min.	15.2	17.3	2.1	0.6 - 3.6	13.8%
Visit time > scheduled ^a	-1.2	1.2	2.3	1.1 - 3.6	191.7%

^aCalculated by subtracting scheduled visit time from observed visit time, among visits for which scheduled duration available.

Figure S3. Adjusted observed time per visit by patient and physician gender

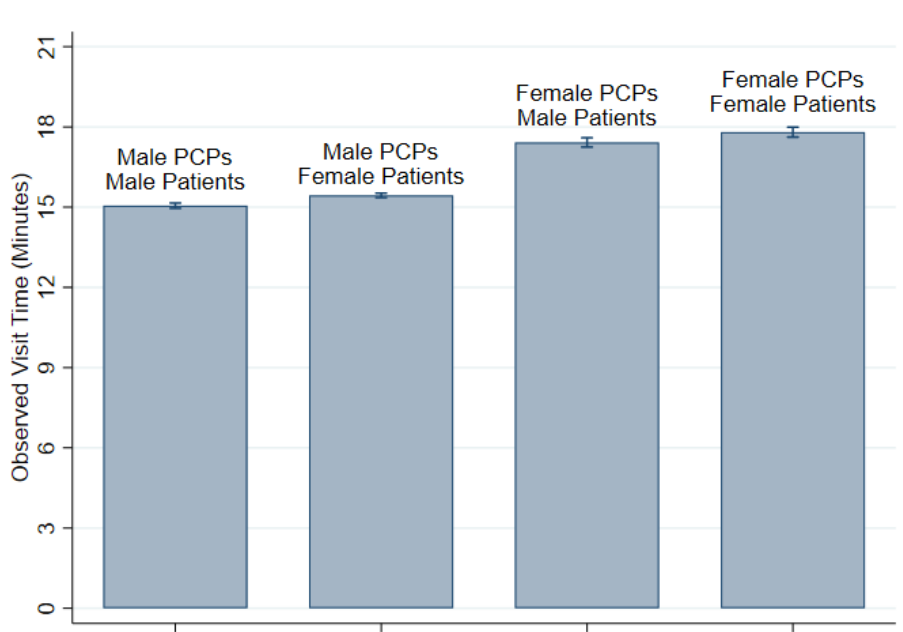


Table S8. Adjusted visit-level outcomes among patients with 2 or more chronic conditions, by physician gender

	Visits to Male PCPs (N=5,003,285)	Visits to Female PCPs (N=2,013,883)	Difference	95% CI for Difference	Difference, %
Allowed charges, \$	122.2	122.9	0.7	-1.7 - 3.0	0.6%
Diagnoses documented, no.	5.5	5.6	0.1	0.0 - 0.2	1.8%
Orders placed, no.	3.8	4.4	0.5	0.4 - 0.7	13.2%
Observed visit time, min.	16.7	19.3	2.7	2.3 - 3.0	16.2%
Visit time > scheduled ^a	0.1	2.2	2.1	1.8 - 2.4	2100.0%

This analysis examined the subset of visits in which two or more chronic conditions were documented based on the CCW algorithm. ^aCalculated by subtracting scheduled visit time from observed visit time, among visits for which scheduled duration available.

Table S9. Adjusted visit-level outcomes among same-day visits, by physician gender

	Visits to Male PCPs (N=3,627,209)	Visits to Female PCPs (N=1,890,493)	Difference	95% CI for Difference	Difference, %
Allowed charges, \$	102.7	101.2	-1.5	-2.9 - -0.1	-1.5%
Diagnoses documented, no.	2.7	2.8	0.1	0.1 - 0.2	3.7%
Orders placed, no.	2.4	2.7	0.3	0.2 - 0.3	12.5%
Observed visit time, min.	14.1	15.8	1.7	1.4 - 2.0	12.1%
Visit time > scheduled ^a	-1.3	-0.1	1.2	0.9 - 1.5	92.3%

This analysis examined the subset of visits that were scheduled as “same-day” visits. ^aCalculated by subtracting scheduled visit time from observed visit time, among visits for which scheduled duration available.

Table S10. Adjusted visit-level outcomes by visit type and physician gender

Visit type	Visits to Male PCPs	Visits to Female PCPs	Difference	95% CI for Difference	Difference, %
<u>Level 3 visits</u>	(N=4,567,873)	(N=2,171,229)			
Allowed charges, \$	94.0	95.2	1.2	-0.5 - 2.9	1.3%
Diagnoses documented, no.	2.9	3.0	0.2	0.1 - 0.2	6.9%
Orders placed, no.	2.1	2.4	0.3	0.2 - 0.4	14.3%
Observed visit time, min.	13.5	15.3	1.7	1.4 - 2.0	12.6%
Visit duration > 25 min., %	11.1	14.6	3.5	2.9 - 4.1	31.5%
Visit time > scheduled ^a	-2.2	-1.0	1.2	0.9 - 1.4	54.5%
<u>Level 4 visits</u>	(N=5,432,695)	(N=2,668,581)			
Allowed charges, \$	128.0	128.9	0.8	-0.4 - 2.1	0.6%
Diagnoses documented, no.	4.7	4.9	0.2	0.1 - 0.2	4.3%
Orders placed, no.	3.6	4.1	0.5	0.4 - 0.6	13.9%
Observed visit time, min.	16.9	19.4	2.5	2.1 - 2.8	14.8%
Visit duration > 40 min., %	4.2	6.1	1.8	1.5 - 2.2	42.9%
Visit time > scheduled ^a	0.5	2.4	1.9	1.6 - 2.2	380.0%
<u>Level 5 visits</u>	(N=314,183)	(N=141,162)			
Allowed charges, \$	165.3	163.3	-2.0	-7.5 - 3.5	-1.2%
Diagnoses documented, no.	6.0	6.0	0.1	-0.1 - 0.2	1.7%
Orders placed, no.	5.0	5.5	0.5	0.2 - 0.8	10.0%
Observed visit time, min.	22.6	25.2	2.6	1.6 - 3.6	11.5%
Visit time > scheduled ^a	3.7	6.6	2.9	1.6 - 4.1	78.4%
<u>New patient visits</u>	(N=803,347)	(N=416,261)			
Allowed charges, \$	168.3	166.0	-2.3	-4.0 - -0.5	-1.4%
Diagnoses documented, no.	3.7	3.9	0.3	0.2 - 0.4	8.1%
Orders placed, no.	3.7	4.6	0.9	0.6 - 1.2	24.3%
Observed visit time, min.	20.0	23.7	3.6	3.0 - 4.3	18.0%
Visit time > scheduled ^a	1.1	3.6	2.5	1.8 - 3.2	227.3%
<u>Preventive visits</u>	(N=1,974,399)	(N=1,335,400)			
Allowed charges, \$	139.9	141.8	2.0	-0.4 - 4.4	1.4%
Diagnoses documented, no.	4.1	4.4	0.3	0.2 - 0.4	7.3%
Orders placed, no.	4.0	4.3	0.3	0.2 - 0.5	7.5%
Observed visit time, min.	18.6	20.6	2.0	1.5 - 2.5	10.8%
Visit time > scheduled ^a	0.9	2.3	1.4	0.9 - 1.8	155.6%

This analysis examined subsets of visits based on their billing codes. Level 3 refers to CPT code 99213, level 4 to CPT code 99214, and level 5 to CPT code 99215. New patient visits included CPT codes 99203-5. Preventive visits included CPT codes 99381-99397, G0402, and G0438-9. ^aCalculated by subtracting scheduled visit time from observed visit time, among visits for which scheduled duration available.

References

1. Neprash HT, Barnett ML. Association of Primary Care Clinic Appointment Time With Opioid Prescribing. *JAMA Netw open*. 2019;2(8):e1910373. doi:10.1001/jamanetworkopen.2019.10373
2. Neprash HT, Zink A, Gray J, Hempstead K. Physicians' participation in medicaid increased only slightly following expansion. *Health Aff*. 2018;37(7):1087-1091. doi:10.1377/hlthaff.2017.1085
3. Barnett ML, Gray J, Zink A, Jena AB. Coupling policymaking with evaluation—the case of the opioid crisis. *N Engl J Med*. 2017;377(24):2306-2309. doi:10.1056/NEJMp1710014
4. García MC, Heilig CM, Lee SH, et al. Opioid prescribing rates in nonmetropolitan and metropolitan counties among primary care providers using an electronic health record system — United States, 2014–2017. *Morb Mortal Wkly Rep*. 2019;68(2):25-30. doi:10.15585/MMWR.MM6802A1
5. Baltrusaitis K, Brownstein JS, Scarpino S V., et al. Comparison of crowd-sourced, electronic health records based, and traditional health-care based influenza-tracking systems at multiple spatial resolutions in the United States of America. *BMC Infect Dis*. 2018;18(1). doi:10.1186/s12879-018-3322-3
6. Lu FS, Hou S, Baltrusaitis K, et al. Accurate influenza monitoring and forecasting using novel internet data streams:a case study in the boston metropolis. *J Med Internet Res*. 2018;20(1). doi:10.2196/publichealth.8950
7. Condition Categories - Chronic Conditions Data Warehouse. <https://www2.ccwdata.org/web/guest/condition-categories>. Accessed January 20, 2020.
8. Ganguli I, Shi Z, Orav EJ, Rao A, Ray KN, Mehrotra A. Declining Use of Primary Care Among Commercially Insured Adults in the United States, 2008–2016. *Ann Intern Med*. 2020;172(4):240-247. doi:10.7326/m19-1834
9. Reid RO, Ashwood JS, Friedberg MW, Weber ES, Setodji CM, Mehrotra A. Retail clinic visits and receipt of primary care. *J Gen Intern Med*. 2013;28(4):504-512. doi:10.1007/s11606-012-2243-x
10. Ware JH, Harrington D, Hunter DJ, D'Agostino RB. Missing data. *N Engl J Med*. 2012;367(14):1353-1354. doi:10.1056/NEJMs1210043
11. Archived Datasets | Data.Medicare.gov | Data.Medicare.gov. <https://data.medicare.gov/data/archives/physician-compare>. Accessed January 20, 2020.