Designing Incentives for Multitasking Agents: Evidence from Payments to English Physicians

Filippo Paternollo (Columbia)
Pietro Tebaldi (Columbia)
Andre Veiga (Imperial)

Bristol, April 2024

Incentive Design in Principal-Agent Problems

- Central to healthcare, education, org econ, etc
- Often, actions and outcomes are high-dimensional
  - doctor chooses tests, prescriptions $\rightarrow$ clinical outcomes
  - teacher chooses topics and methods $\rightarrow$ test scores, human capital, etc
- Often, there is multitasking: higher effort in one task influences the cost of other tasks [Holmstrom and Milgrom, 1991]

Empirical models of multitasking

- Counterfactuals require estimates of
  - distribution of types
  - interaction between outcomes
- Each task potentially interacts with all other tasks
  - number of parameters grows quadratically
- Most applied work focuses on testing for multi-tasking

This Paper

- Empirically tractable model of multitasking
- Sufficient conditions for identification
- Application to Quality of Outcomes Framework (QOF) in England 2009-2019
  - world's largest P4P scheme in primary care
- Strong evidence of
  - physicians responding to financial incentives
  - interactions between indicators
- Variation:
  - in QOF incentives over time
  - practice location exogenously shifts patient composition
- Estimate model & counterfactual design of incentives
  - QOF increases average achievement by $\approx 40\%$
  - optimal re-design increases payer's utility by 3%

Literature

- Empirical models of multitasking: Slade [1996], Buser and Peter [2012], Hong, Hossain, List, and Tanaka [2018], Goes, Ilk, Lin, and Zhao [2018], Manthei and Sliwka [2019], Rodríguez-Lesmes and Vera-Hernández [2021], Kim, Sudhir, and Uetake [2022], Dinerstein and Opper [2022]
  - We go beyond testing
  - We can consider counterfactual design
- Pay-for-performance in healthcare: Gaynor et al. [2004], Dumont et al. [2008], Mullen et al. [2010], Choné and Ma [2011], Clemens and Gottlieb [2014], Li et al. [2014], Einav et al. [2018], Gupta [2021], Rodríguez-Lesmes and Vera-Hernández [2021], Einav et al. [2022], Gaynor et al. [2023], Dunn et al. [2024], Shi [2024, and many more]
  - We incorporate multitasking
  - We focus on primary care in non-US context

Roadmap

GP practices (GPPs)

- GPPs provide primary care: prescriptions, minor interventions, referral to secondary care
- Approximately 8000 GPPs in England
- Each GPP has about 5 doctors (but we study GPPs)
- Zero prices to patients
- Revenue:
  - $\approx$ 75% capitation (# of individuals registered, mild risk adjustment)
  - $\approx$ 25% financial incentives, mainly from QOF

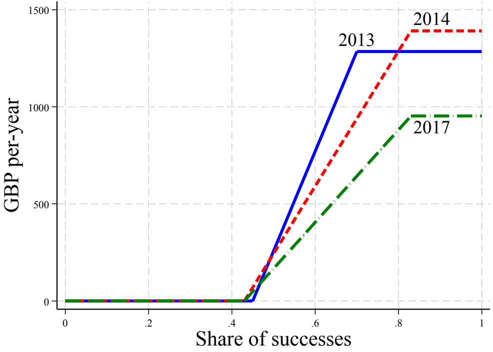
QOF

- Gives GPPs yearly financial incentives to perform certain tasks (“indicators”):
  - “The percentage of patients aged 75 or over with a fragility fracture on or after 1 April 2012, who are currently treated with an appropriate bone-sparing agent”
  - “The percentage of patients on the chronic kidney disease (CKD) register in whom the last blood pressure reading, measured in the previous 15 months, is 140/85 or less.”
- Started 2004, changes over time in
  - intensity of incentives
  - which tasks are incentivized
- Voluntary participation (95.1% in 2019)
- Total payments $\approx$ £1B
- Electronic record-keeping $\rightarrow$ minimal errors / cheating
- We focus on 40 indicators that are “truly clinical”

QOF payments

- Indicator j has n_j relevant patients
- If task is successful for k_j patients, achievement is $k_j/n_j = y_j \in [0,1]$
- Revenue per patient for indicator j is $\rho_{jt}(y_{ijt})$: piece-wise linear with slope α_{jt} and thresholds $\underline{y_{jt}}, \overline{y_{jt}}$:

DM11: % diabetes patients in whom the last glycohaemoglobin IFCC-HbA1c is 64 mmol/mol or less



Data

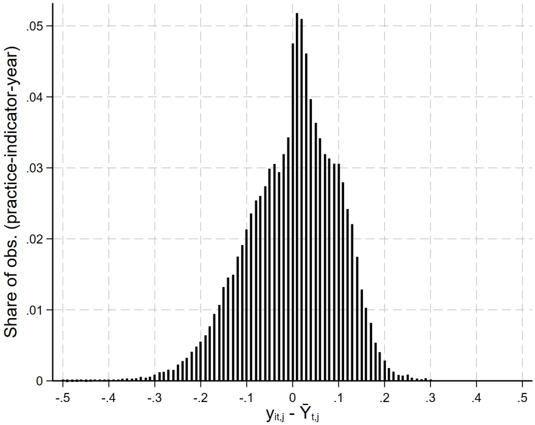
- GPP i, indicator j, year t
- Achievement y_{ijt}
- GPP covariates x_{it} (number of physicians, average age, etc)
- Nr of relevant patients n_{ijt}
- Thresholds $\overline{y_{jt}}, \underline{y_{jt}}$
- Incentives α_{jt}
- Everything in 2020 £

Summary Stats

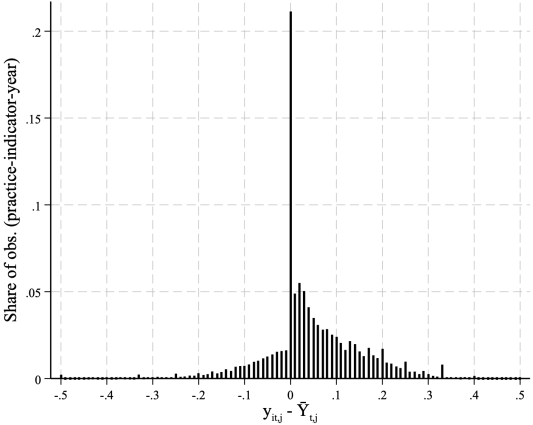
		Mean	Std. Dev.	P-10	Median	P-90	Obs.
	Panel B:	**Indicator-practice-year**					
Share of successes	$y_{it,j}$	0.878	0.121	0.722	0.906	1	2389540
Number of patients	$n_{it,j}$	177	264	4	87	445	2389540
Revenues per patient	$\rho_j(y_{it,j})$	11.16	12.64	2.32	7.23	27.99	2389540

Practices respond to incentives

DM11: % diabetes patients in whom the last glycohaemoglobin IFCC-HbA1c is 64 mmol/mol or less



For all indicators, distribution of $y_{ijt} - \overline{y_{jt}}$



- Bunching suggests strong response to financial incentives
- $y_{ijt} > \overline{y_{jt}}$ suggests GPPs also have non-financial incentives and/or some tasks are complements.

Achievement responds to incentives

- Column (4) uses variation in incentives within GPP-indicator over time

Panel B: Share of successes – $y_{it,j} = \delta^1 \overline{Y}_{t,j} + \delta^2 \alpha_{t,j} + \zeta^1 x_i + \zeta^2 \widehat{\xi}_i + \epsilon_{it,j}$					
$\overline{Y}_{t,j}$	0.281	0.276	0.241	0.279	0.289
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\overline{Y}_{t,j}$ × Large practice				0.076	0.004
				(0.002)	(0.000)
$\alpha_{t,j}$	0.087	0.082	0.031	0.094	0.170
	(0.005)	(0.005)	(0.007)	(0.005)	(0.006)
$\alpha_{t,j}$ × Large practice				0.115	−0.135
				(0.009)	(0.004)
Practice-level average residual $\hat{\xi}_i$ from practice-choice model					0.129
					(0.005)
FE	Ind.	Ind., Practice	Ind.× Practice	Ind.× Practice	Ind.
Controls	N	N	N	N	Y
R-squared	0.412	0.476	0.656	0.656	0.431
Observations	2353922	2353922	2332413	2332413	2060431

Roadmap

2 tasks

- Single GPP, 2 tasks ($j = 1, 2$)
- Number of patients n_j
- A clinic chooses outcomes $y_j \in [0, 1]$
  - large number of patients $\rightarrow$ negligible outcome noise
- Payer chooses revenue functions $\rho_j(\cdot)$
- GPP utility:

$$
\begin{aligned}
U(y) = & n_1 \rho_1(y_1) + n_2 \rho_2(y_2) + n_1 \theta_1 y_1 + n_2 \theta_2 y_2 \\
& - n_1 \lambda_1 y_1^2 - n_2 \lambda_2 y_2^2 - (n_1 + n_2) \lambda_{12} y_1 y_2
\end{aligned}
$$

- Revenue, Altruism, Costs, Complementarities
  - (Altruism is short-hand for all non-financial incentives)
- If $\lambda_{12} > 0$, tasks are "substitutes"
- If $\lambda_{12} < 0$, tasks are "complements"

Many tasks

$$U(y) = \sum_j n_j \left(\rho_j(y_j) + \theta_j y_j\right) - y^T \Lambda y$$

$$\Lambda = \begin{bmatrix} n_1\lambda_1 & n_2\lambda_{12} & \cdots & n_J\lambda_{1J} \\ n_1\lambda_{12} & n_2\lambda_2 & & \\ \vdots & & \ddots & \\ n_1\lambda_{1J} & n_2\lambda_{2J} & & n_J\lambda_J \end{bmatrix}$$

- For each j, the n_j patients are identical
- If all n_j scale up, solution is unchanged
- If n_j increases, other things equal, incentives for task j increase

Roadmap

1. Setting and Data
2. Model
3. **Demand**
4. Identification & Estimation
5. Estimates & GOF
6. Counterfactuals
7. Conclusion

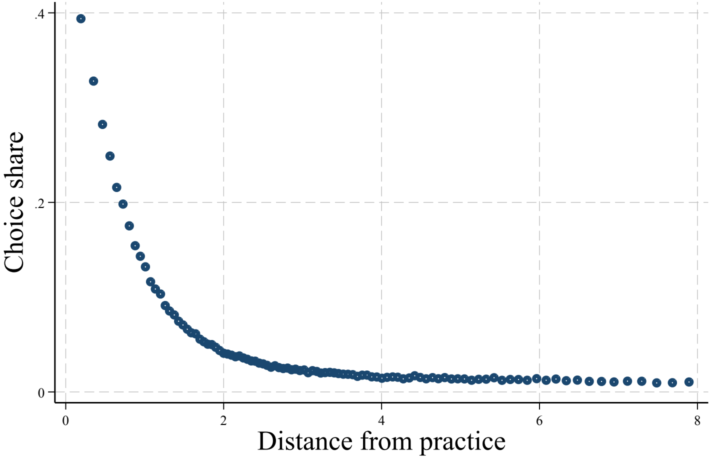
Demand

- Patients might select into "high-quality" practices
  - must consider demand
- We use exogenous variation in patient-practice location z_{it} to identify choice-relevant unobservable quality
- Simple logit: share of patients from location ℓ choosing practice i in year t is

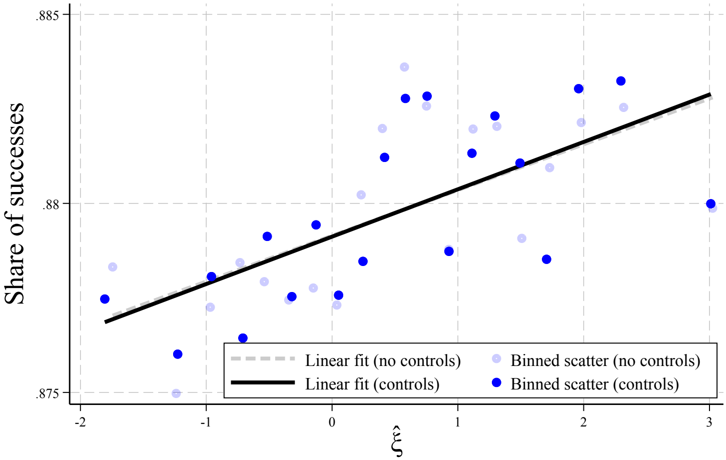
$$s_{it}^{\ell} = \frac{\exp\left\{\gamma z_i^{\ell} + \eta x_i + \xi_{it}^{\ell}\right\}}{1 + \sum_{d: z_d^{\ell} \leq 5} \exp\left\{\gamma z_d^{\ell} + \eta x_d + \xi_{dt}^{\ell}\right\}} \qquad , z_i^{\ell} \leq 5 \text{ kilometers.}$$

- Aggregate $\widehat{\xi}_i = \mathbb{E}_{\ell,t}[\xi_{it}^{\ell}]$
- **Assumption:** n_{ijt} iid conditional on $x_i, \widehat{\xi}_i$
- Currently working on adding health conditions by location (i.e., estimate ξ_{ij})
  - we will approximate market size for GPP i condition j: total number of patients with that condition in all GPPs within 10Km of i
  - allows patients with condition j to choose GPP i because it is high-quality for condition j (but possibly low-quality for $k \neq j$)

Distance shifts demand



Demand residual is correlated with achievement



Roadmap

Identification

- Goal: identify technology Λ separately from type distribution $f(\theta)$
- Data: demand n_{it}, achievement y_{it}, characteristics x_{it}, z_{it}, incentives $\rho_{jt}(\cdot)$
- We prove that the model is identified if:
- **Assumption (LQU):** Utility $U(y)$ is Linear-Quadratic
- **Assumption (exogeneity):** Instrument z_{it} such that demand is $n_{it} = \sigma(x_{it}, z_{it}, \xi_{it})$
  - n_{it} and θ_{it} independent conditional on (x_{it}, ξ_{it})
- **Assumption (independence):** $f(\theta_{it}|x_{it}, \xi_{it}) = \prod_j f_j(\theta_{it,j}|x_{it}, \xi_{it})$
  - currently working on relaxing.
- **Assumption (variation):** Rich variation in incentives (α_{jt}) and task assignments (z_{it}) to distinguish any function of y_{it}
  - Intuition: Change incentives for blood sugar control. Compare GPP A with many diabetics to GPP B with few. If A's cholesterol outcomes improve more than B's $\rightarrow$ cholesterol & blood sugar are complements

Estimation

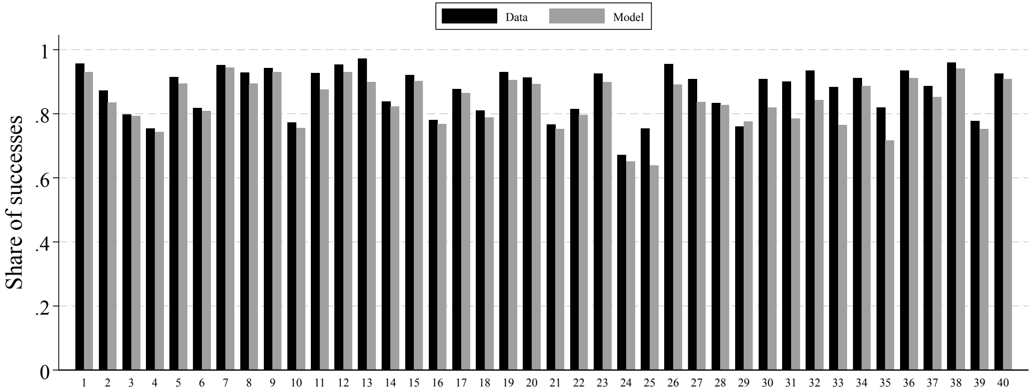
- Assume y_{ijt} is the optimal effort choice made by the GPPs (up to integers)
- Let $\tilde{x}_i = \left(x_i, \hat{\xi}_i\right)$ be GPP covariates, including demand residuals
- Assume that, for each ijt,

$$\theta_{ijt} \sim \mathcal{N}\left(\mu_j \tilde{x}_i, \sigma_j\right), \qquad (n_{it} \mid \tilde{x}_i) \perp (\theta_{it} \mid \tilde{x}_i)$$

- Given LQU, $\frac{\partial U_{it}}{\partial y_{ijt}}$ is linear in θ_{ijt}
  - Can derive (discrete-continuous) distribution of θ_{ijt} analytically: ▸ Details
- Estimate Λ and $\{\sigma_j, \mu_j\}_j$ by MLE (1060 parameters)

Roadmap

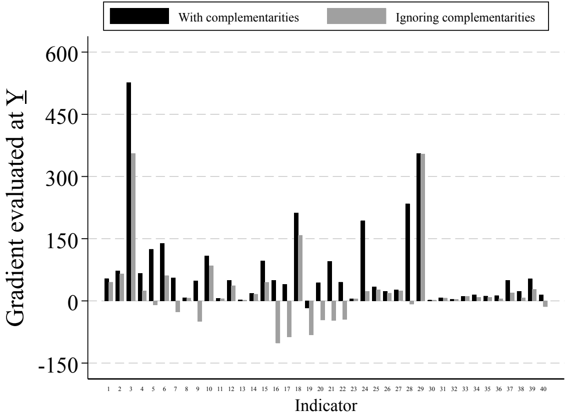
Goodness of Fit



1	In those patients with Atrial Fibrillation in whom re is a record of a CHADS2 score of 1 (latest in preceding 15 months), % of patients who are ly treated with anti-coagulation drug rapy or an anti-platelet rapy.
2	In those patients with Atrial Fibrillation whose latest record of a CHADS2 score is greater than 1, % of patients who are ly treated with anti-coagulation drug rapy	3.03
3	% of patients with hypertension in whom last blood pressure (measured in previous 9 months) is 150/90 or less.	0.21	0.00
4	% of patients with coronary heart disease who are ly treated with a beta blocker (unless a contraindication or side -effects are recorded).	1.36	-0.10	-1.52
5	% of patients with coronary heart disease in whom last blood pressure reading (measured in previous 15 months) is 150/90 or less.	-0.17	0.51	-7.48	-6.10
6	% of patients with coronary heart disease whose last measured total cholesterol (measured in previous 15 months) is 5mmol/l or less.	-1.94	0.21	-4.14	-6.43	-7.47
7	% of patients with coronary heart disease with a record in preceding 12 months that aspirin, an alternative anti-platelet rapy, or an anti-coagulant is being taken	0.98	0.53	-0.78	-5.70	-15.66	-4.26
8	% of patients with a history of myocardial infarction (from 1 April 2011) ly treated with an ACE inhibitor (or ARB if ACE intolerant), aspirin or an alternative anti-platelet rapy, beta-blocker and statin	0.45	0.87	-0.19	0.10	-0.43	0.26	-0.92
9	% of patients with coronary heart disease who have had influenza immunisation in preceding 1 September to 31 March	-0.65	-1.87	-0.53	-2.55	-16.11	-5.89	-13.72	1.21
10	% of patients on CKD register in whom last blood pressure reading, measured in previous 15 months, is 140/85 or less.	-1.86	0.63	1.85	-2.49	-1.75	-3.31	0.92	0.12	-1.17
11	% of patients on CKD register with hypertension and proteinuria who are treat ed with an angiotensin converting enzyme inhibitor (ACE-I) or angiotensin receptor blocker (ARB) (unless a contraindication or side effects are recorded).	0.18	-0.18	-0.21	0.88	-0.71	1.30	-0.98	-1.02	1.08	0.00
12	% of patients with COPD who have had influenza immunisation in preceding 1 September to 31 March.	-0.78	-0.14	0.54	-1.29	-1.74	1.00	0.86	0.06	-3.33	-0.20	0.10
13	In those patients with a new diagnosis of hypertension aged 30 or over and who have not attained age of 75, recorded between preceding 1 April to 31 March, who have a recorded CVD risk assessment score of =20% in preceding 12 months: % who are ly treated with statins	0.29	0.35			0.36	-0.83	-1.12	2.69	0.77	0.05	-2.62	0.41
14	% of patients aged 18 or over with a new diagnosis of depression in preceding 1 April to 31 March, who have been reviewed not earlier than 10 days after and not later than 56 days after date of diagnosis	-0.24	0.72			0.78	-1.66	-0.12	-0.39	0.52	0.04	-0.10	0.17	-0.54
15	% of patients with diabetes in whom last HbA1c is 10 or less (or equivalent test/reference range depending on local laboratory) in previous 15 months.	1.74		-5.96	1.01	0.08	-1.49	0.55	0.48	2.22	1.42	0.85	3.70		
16	% of patients with diabetes, on register, in whom last IFCC-HbA1c is 64 mmol/mol or less in preceding 12 months	1.00	1.23	-4.58	0.29	0.87	-0.42	0.62	0.43	0.82	-1.00	1.33	1.12	1.26	-0.50	
17	% of patients with diabetes in whom last HbA1c is 9 or less (or equivalent test/reference range depending on local laboratory) in previous 15 months.	0.87	1.07	-3.71	0.73	0.35	0.03	-0.36	-0.02	1.36	1.40	-0.07	2.26	-1.65	-0.10		-28.3
18	% of patients with diabetes in whom last blood pressure is 145/85 or less.	0.36		-6.91	-0.37	-3.06	1.47	-0.25	-0.44	5.09	-2.75	-0.14	0.31			-0.86	0.71	-2.11
19	% of patients with diabetes who have had influenza immunisation in preceding 1 September to 31 March.	-0.04	0.30	-0.15	-0.63	2.76	2.93	2.43	-0.42	-11.60	0.78	-0.69	-1.90	-0.16	-0.27	-2.65	-0.19	-2.88	-5.49
20	% of patients with diabetes in whom last blood pressure is 150/90 or less.	0.43	0.39	-3.59	-2.03	-1.07	1.31	1.32	0.40	2.91	0.12	-0.01	1.03	1.04	-0.03		1.99	-1.26		-4.06
21	% of patients with diabetes, on register, in whom last blood pressure reading (measured in preceding 12 months) is 140/80 mmHg or less	0.63	-0.54	-2.53	0.45	-3.25	3.06	-2.19	-0.22	1.74	-3.82	0.12	0.01	-1.25	-0.06		-1.05	0.31		-0.91	-21.00
22	% of patients with diabetes whose last measured total cholesterol within preceding 15 months is 5mmol/l or less	-1.09	-0.98	-0.35	-0.70	1.28	-4.55	-1.70	0.27	3.69	0.19	-0.01	-0.61	0.07	0.36	-7.33	-2.98	-5.63	-7.57	-5.26	-1.60	-5.01
23	% of patients with diabetes, on register, with a diagnosis of nephropathy (clinical proteinuria) or micro-albuminuria who are ly treated with an ACE-I (or ARBs)	0.35	0.13	-0.19	0.13	0.46	-0.30	-0.25	0.17	0.15	0.06	-0.59	0.06	0.16	-0.17	0.44	0.09	0.35	0.02	-0.42	0.16	-0.13	0.02
24	% of patients with diabetes in whom last HbA1c is 7 or less (or equivalent test/reference range depending on local laboratory) in previous 15 months.	-2.18	-1.26	-5.91	1.05	-0.91	0.12	-2.28	-0.36	0.34	-1.06	-0.51	-2.10	0.81	0.84	-10.27	-17.03	-8.18	3.29	-1.09	-6.27	-1.58	-7.50	-0.06
25	% of patients age 18 and over on drug treatment for epilepsy who have been seizure free for last 12 months recorded in previous 15 months.	-4.84	0.30	-0.30	1.83	-0.29	-0.57	0.17	0.22	-0.26	-3.81	-0.88	-1.87	0.80	0.53	0.30	-0.85	1.15	0.90	-1.00	2.73	-1.64	0.52	0.50	-0.68
26	% of patients with a diagnosis of heart failure due to Left Ventricular Dysfunction (LVD) who are ly treated with an ACE inhibitor or Angiotensin Receptor Blocker (ARB) , who can tolerate rapy with no contra -indication.	-1.44	-0.37	-0.57	1.52	0.63	-3.36	-1.99	-0.34	0.26	0.08	0.01	-0.50	0.27	0.14	-0.24	-0.51	2.15	-0.61	-0.85	-0.89	0.11	-0.31	-0.66	1.57	-0.61
27	% of patients with heart failure due to LVD who are ly treated with an ACE inhibitor or Angiotensin Receptor Blocker, who are additionally treated with a beta -blocker licensed for heart failure, or recorded as intolerant	-1.11	-1.44	0.02	-0.86	-0.06	0.14	0.68	-0.29	0.80	0.12	-1.72	-0.76	-0.81	-0.38		0.02	0.36	0.45	-0.49	0.24	-0.46	-0.95	0.31	1.07	1.45	-12.40
28	% of patients with hypertension in whom last blood pressure reading (measured in preceding 9 months) is 150/90 mmHg or less	-0.37	-0.39			-8.55	0.06	-2.21	-0.21	-2.38	-1.50	-0.06	-0.46	-0.20	-0.24		-4.56	-2.96		0.48	-1.77	-3.19	-1.99	-0.18	-7.21	0.74	-0.71	-0.03
29	% of patients aged 79 or under with hypertension in whom last blood pressure reading (measured in preceding 9 months) is 140/90 mmHg or less	0.97	0.04			3.37	-4.89	1.72	0.05	2.37	2.54	-0.14	1.55	0.03	0.08		1.15	0.57		-0.73	-1.76	-0.67	2.87	0.01	0.69	-2.17	1.05	-0.14	-3.79
30	% of patients on lithium rapy with a record of lithium levels in rapeutic range within previous 6 months.	-0.27	0.08	-0.18	0.74	-0.02	-0.32	0.85	-0.95	-0.73	0.20	-1.71	0.17	0.00	-0.35	-0.59	-0.39	-0.22	0.24	0.01	0.10	-0.09	0.08	0.03	0.76	-1.05	-1.41	-0.02	-0.21	0.26
31	% of women with schizophrenia, bipolar affective disorder and or psychoses whose notes record that a cervical screening test has been preformed in preceding 5 years.	-0.46	-1.43	-0.20	0.52	0.83	-0.70	0.80	0.92	0.02	-0.25	-1.43	-0.92				1.08	-0.30		-0.54	-1.66	0.55	1.00	0.05	-0.42	0.70	-0.71	1.84			0.49
32	% of patients aged between 50-74, with a fragility fracture, in whom osteoporosis is confirmed on DXA scan, who are ly treated with an appropriate bone-sparing	1.28	0.73	-0.38	-1.61	-0.19	1.21	-0.31	0.55	0.05	-0.24	2.67	-0.66	0.58	0.14		0.59	-0.49		0.13	0.40	0.22	-0.49	-0.60	-0.09	1.70	0.11	-1.15	-0.14	-0.67	-1.25	-0.14
33	% of patients aged 75 or over with a fragility fracture on or after 1 April 2012, who are ly treated with an appropriate bone-sparing agent	-0.54	-1.14	-0.66	0.51	-0.73	0.38	0.07	-1.26	0.30	0.33	-0.58	0.68	0.17	-0.56		1.37	0.64		-0.28	-1.20	0.79	0.87	-0.40	-2.27	-0.25	2.46	-0.86	-0.16	-0.22	-0.33	2.84	-0.98
34	% of patients with peripheral arterial disease in whom last blood pressure reading (measured in preceding 15 months) is 150/90 or less	3.10	-0.16	-0.28	0.22	-2.53	0.81	1.84	0.58	0.14	-0.63	-0.27	0.18	-0.03	-0.24		1.48	1.68		0.17	0.19	-1.86	-0.39	0.03	-1.30	-0.41	-4.86	-1.09	-0.22	-0.03	-0.15	-0.11	-0.06	-1.42
35	% of patients with peripheral arterial disease in whom last measured total cholesterol (measured in preceding 15 months) is 5.0mmol/l or less	3.48	-0.94	0.41	2.57	-0.29	-4.30	0.66	-0.82	0.53	-1.06	0.56	-1.33	0.02	0.38		0.00	-0.96		0.21	0.42	0.20	-0.01	0.73	-0.38	-0.62	5.24	1.45	2.52	-2.36	-0.46	0.58	3.25	-2.20	-2.62
36	% of patients with peripheral arterial disease with a record in preceding 15 months that aspirin or an alternative anti-platelet is being taken	0.49	-0.10	0.06	-0.97	1.46	1.05	-2.32	0.16	0.96	-0.39	0.30	-1.68	0.18	-0.76		0.62	0.99		-0.55	0.68	-0.68	-0.17	-0.09	-1.21	0.56	-1.84	-0.90	-0.13	-0.05	0.17	0.18	0.47	-0.14	-12.56	-1.34
37	% of patients with a history of TIA or stroke in whom last blood pressure reading (measured in previous 15 months) is 150/90 or less.	-1.55	-0.53	-0.29	0.86	-8.41	1.06	2.30	0.58	3.87	-0.29	0.85	-0.77	-0.53	-0.85	0.66	0.25	0.64	-1.28	0.52	-0.19	-0.35	0.22	-0.14	-0.92	1.07	0.78	-0.09	-0.53	0.46	-0.40	-1.27	0.32	-0.41	-0.68	-0.36	0.47
38	% of patients with a stroke shown to be non-haemorrhagic, or a history of TIA, who have a record that an anti-platelet agent,	-1.77	-4.77	0.01	-0.21	1.89	2.40	-3.86	0.41	1.69	-0.07	-0.23	-0.95	0.18	-0.91	1.08	1.07	0.38	-0.61	0.47	0.48	-0.87	-0.48	0.19	-1.01	-3.96	2.28	-0.61	0.16	0.35	0.28	0.57	0.43	0.42	1.79	1.30	-0.91	-5.10
39	% of patients with TIA or stroke whose last measured total cholesterol (measured in preceding 15 months) is 5 mmol/l or less	-3.15	1.55	0.16	-1.58	1.54	-8.90	1.94	-0.75	1.90	-3.48	-2.36	-1.98	1.02	5.50	0.63	0.07	0.58	2.77	-0.59	0.66	2.23	-2.69	0.05	0.73	2.60	0.81	-0.53	3.19	-3.95	-0.29	0.23	-0.41	3.32	-0.46	-4.32	-1.38	-5.84	-4.45
40	% of patients with TIA or stroke who have had influenza immunisation in preceding 1 September to 31 March.	0.55	-1.13	-0.48	1.65	3.61	1.37	2.77	-0.47	-10.68	-0.21	0.57	-1.57	0.55	-0.50	0.90	0.20	0.63	0.95	-1.80	0.59	0.43	0.71	0.00	-0.29	1.41	1.94	0.50	-0.50	0.43	0.33	0.57	0.26	1.80	-0.58	-0.21	-0.10	-13.61	-8.44	-5.30

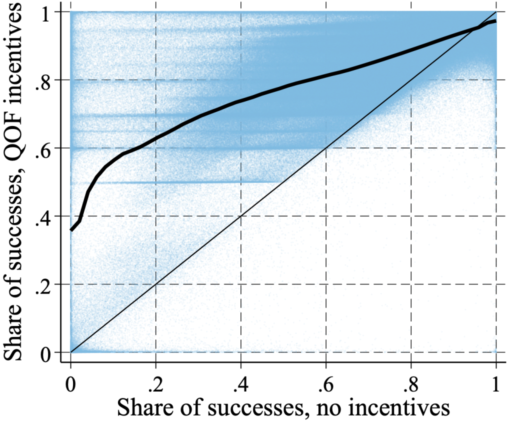
How important is Multitasking?

- Given estimates, for each j, calculate $\mathbb{E}_{it}[\partial U_{it}/\partial y_{it,j}]$, evaluated at $\underline{y_{tj}}$ (beginning of incentives)
    - then repeat this using $\lambda[j,k] = 0$ if $j \neq k$

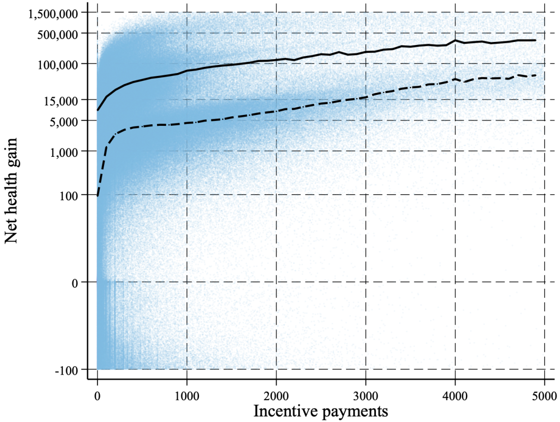


Roadmap

Shutting Down QOF: achievement



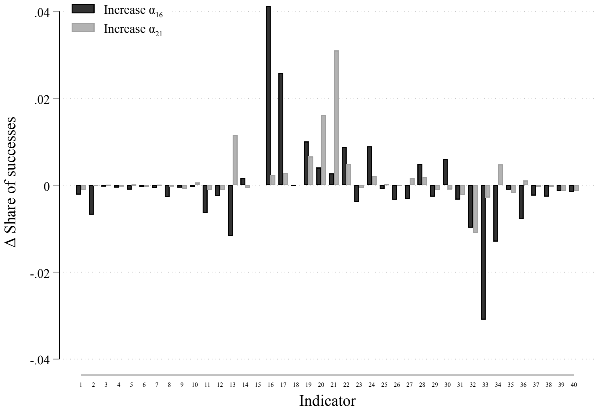
Shutting Down QOF: health gains (in £)



- Ratio of QOF payments to median health gains (in £) is ≈1:5

Increasing incentives for one indicator

- Increase incentives for indicators16 & 21 (by £40)
- Significant impact on other indicators



Empirical incentive design

- b_j are health benefits net of medical costs for indicator j (in £, from NICE guidelines)
  - so far only 20/40 indicators
- Set $\underline{y_j}$ fixed and $\overline{y_j} = 1$ for each years
- Choose α_j to maximize the payer's objective

$$W = \sum_{i,j,t} n_{ijt} \int \left(y_{ijt} b_j - \rho_{jt} \left(y_{ijt} \mid \alpha_{jt} \right) \right) f \left(\theta_{ijt} \mid \tilde{x}_i \right) d\theta_{ijt}$$

where y_{ijt} is chosen by GPPs to maximize utility

- Requires, for each α, solving the problem for all GPPs $\rightarrow$ unfeasible
- We use k-means to cluster GPPs in terms of x_i, ξ_i, n_{ijt}
  - For every group $g = 1, \ldots, 20$ obtain weight π_g and average values x_g, ξ_g, n_{gjt}
  - Maximize this approximate W. At the solution, compute outcomes for all GPPs

Optimal incentives increase payer utility by 3%

	No QOF Δ from QOF	QOF	Optimized QOF Δ from QOF
Practice payoffs	-348 -11%	3,240	164 5%
QOF payments	-361 -100%	361	221 61%
Medical costs	-1,449 -3%	43,465	731 2%
Health benefits	-5,574 -4%	131,900	3,915 3%
Welfare	-4,113 -5%	91,314	3,128 3%

Notes: All monetary values are in GBP millions. Welfare is computed as the Practice payoffs + Health benefits - QOF payments - Medical costs

Roadmap

Next Steps

- Include co-morbidities in demand
- Integrate “missing indicators” into the estimation

Conclusion

- Empirically tractable principal-agent model with multitasking
- Sufficient conditions for identification
  - relying on variation in exposure to different tasks
- Apply model to QOF program in England
- Ample evidence of response to incentives and multitasking
- Model allows counterfactuals:
  - Program generates large welfare gains
  - Scope for optimization of incentives accounting for multitasking

Thank you

a.veiga@imperial.ac.uk

Additional slides:

$U'(y)$ Linear in ε

- Given LQU, $\frac{\partial U_{it}}{\partial y_{ijt}}$ is linear in θ_{ijt}
- For instance, in the 2D case:

$$\frac{\partial U}{\partial y_1} = n_1 \rho_1'(y_1) + n_1 \theta_1 - 2 n_1 \lambda_1 y_1 - (n_1 + n_2) \lambda_{12} y_2$$

- If $y_1 = 1$, and knowing $\rho_1'(1) = 0$, then

$$\frac{\partial U}{\partial y_1}|_{y_1=1} \geq 0 \Leftrightarrow \theta_1 \geq 2\lambda_1 + \frac{n_1 + n_2}{n_1} \lambda_{12} y_2$$

- If $y_1 \in (\overline{y_1}, 1)$, the FOC holds, so

$$\frac{\partial U}{\partial y_1} = 0 \Leftrightarrow \theta_1 = 2\lambda_1 y_1 + \frac{n_1 + n_2}{n_1} \lambda_{12} y_2 - \rho_1'(y_1)$$

Back

Thomas Buser and Noemi Peter. Multitasking. *Experimental economics*, 15(4):641–655, 2012.
Philippe Choné and Ching-to Albert Ma. Optimal health care contract under physician agency. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, pages 229–256, 2011.
Jeffrey Clemens and Joshua D Gottlieb. Do physicians' financial incentives affect medical treatment and patient health? *American Economic Review*, 104(4):1320–1349, 2014.
Michael Dinerstein and Isaac M Opper. Screening with multitasking: Theory and empirical evidence from teacher tenure reform. Technical report, National Bureau of Economic Research, 2022.
Etienne Dumont, Bernard Fortin, Nicolas Jacquemet, and Bruce Shearer. Physiciansâ multitasking and incentives: Empirical evidence from a natural experiment. *Journal of health economics*, 27(6):1436–1450, 2008.
Abe Dunn, Joshua D Gottlieb, Adam Hale Shapiro, Daniel J Sonnenstuhl, and Pietro Tebaldi. A denial a day keeps the doctor away. *The Quarterly Journal of Economics*, 139(1):187–233, 2024.
Liran Einav, Amy Finkelstein, and Neale Mahoney. Provider incentives and healthcare costs: Evidence from long-term care hospitals. *Econometrica*, 86(6):2161–2219, 2018.
Liran Einav, Amy Finkelstein, Yunan Ji, and Neale Mahoney. Voluntary regulation: Evidence from medicare payment reform. *The quarterly journal of economics*, 137(1):565–618, 2022.
Martin Gaynor, James B Rebitzer, and Lowell J Taylor. Physician incentives in health maintenance organizations. *Journal of Political Economy*, 112(4):915–931, 2004.
Martin Gaynor, Nirav Mehta, and Seth Richards-Shubik. Optimal contracting with altruistic agents: Medicare payments for dialysis drugs. *American Economic Review*, 113(6):1530–1571, 2023.
Paulo B Goes, Noyan Ilk, Mingfeng Lin, and J Leon Zhao. When more is less: Field evidence on unintended consequences of multitasking. *Management Science*, 64(7):3033–3054, 2018.
Atul Gupta. Impacts of performance pay for hospitals: The readmissions reduction program. *American Economic Review*, 111(4):1241–1283, 2021.
Bengt Holmstrom and Paul Milgrom. Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics, and Organization*, 7(special_issue):24–52, 1991.
Fuhai Hong, Tanjim Hossain, John A List, and Migiwa Tanaka. Testing the theory of multitasking: Evidence from a natural field experiment in chinese factories. *International Economic Review*, 59(2):511–536, 2018.
Minkyung Kim, K Sudhir, and Kosuke Uetake. A structural model of a multitasking salesforce: Incentives, private information, and job design. *Management Science*, 68(6):4602–4630, 2022.
Jinhu Li, Jeremiah Hurley, Philip DeCicca, and Gioia Buckley. Physician response to pay-for-performance: Evidence from a natural experiment. *Health economics*, 23(8):962–978, 2014.
Kathrin Manthei and Dirk Sliwka. Multitasking and subjective performance evaluations: Theory and evidence from a field experiment in a bank. *Management Science*, 65(12):5861–5883, 2019.
Kathleen J Mullen, Richard G Frank, and Meredith B Rosenthal. Can you get what you pay for? pay-for-performance and the quality of healthcare providers. *The Rand journal of economics*, 41(1):64–91, 2010.
Paul Rodríguez-Lesmes and Marcos Vera-Hernández. Multitasking, two-part contracts, and bunching: an application to doctorsâ jobs and incentive contracts. *Mimeo, University College London*, 2021.