Electronic Companion

EC.1. Notation

Table EC.1 summarizes the main notation used in the paper.

Notation (subscript i refers to consumer i)	
h	Denotes an amount of FV spending
u	Denotes an amount of OF spending
$\mathcal{P}$	Population of individuals $i \in \mathcal{P}$ with cardinality P
$B_i(\cdot)$	Total food budget each shopping trip. B_{i_0} denotes consumer i's food budget in the absence of interventions.
$v_i(\cdot)$	Value of nutrition. v_{i_0} denotes consumer i's value of nutrition in the absence of interventions.
$f_{u_i}(\cdot)$	Utility for OF, a function of the amount of money spent on OFs each shopping trip.
$f_{h_i}(\cdot)$	Utility for FV, a function of the amount of money spent on FVs each shopping trip.
$h_i(\cdot)$	Optimal FV spending of consumer i as a function of the interventions.
$u_i(\cdot)$	Optimal OF spending of consumer i as a function of the interventions.
β	Monetary investment in the monetary intervention
α	Monetary investment in the education intervention
$\hat{\alpha}$	Variable that maps onto α and β as: $\alpha \leftarrow \hat{\alpha}$, $\beta \leftarrow 1-\hat{\alpha}$
$\hat{\alpha}_i^*$	Consumer i's individually optimal intervention bundle
U	Planner's budget per person
$\mathcal{M}$	A set of intervention bundles with cardinality M
τ_i	Target FV spending for individual i

Table EC.1 Summary of notation used in the main text.

EC.2. Proofs of Analytical Results

EC.2.1. Proof of Proposition 1 and Corollary 1

Proof of Proposition 1. In the following proof we will drop the subscript i for simplicity. The proof of each of these statements relies on computing the implicit derivatives of $h(\alpha, \beta)$

and $u(\alpha,\beta)$ with respect to α and β, which arise from the Karush-Kuhn-Tucker (KKT) conditions of Problem 1, given by:

$$\begin{cases} (1+v(\alpha))f_h'(h) - f_u'(u) = 0 \\ u + h - B(\beta) = 0. \end{cases} \tag{EC.1}$$

With respect to α (and writing $h(\alpha,\beta)$ as $h(\alpha)$ for simplicity), it can be shown that these implicit derivatives are given by:

$$h'(\alpha) = -\frac{v'(\alpha)f_h'(h(\alpha))}{(v(\alpha)+1)f_h''(h(\alpha)) + f_u''(u(\alpha))} \tag{EC.2}$$

$$u'(\alpha) = \frac{v'(\alpha)f_h'(h(\alpha))}{(v(\alpha)+1)f_h''(h(\alpha)) + f_u''(u(\alpha))} \tag{EC.3}$$

Under the assumption that $f_h(\cdot)$ and $f_u(\cdot)$ are concave increasing (Assumption 1), and that $v'(\alpha)$ is positive (Assumption 2), it is clear that $h'(\alpha)$ is positive while $u'(\alpha)$ is negative.

With respect to β, it can be shown that these implicit derivatives are given by:

$$h'(\beta) = \frac{B'(\beta)f_u''(u(\beta))}{(v(\alpha)+1)f_h''(h(\beta)) + f_u''(u(\beta))} \tag{EC.4}$$

$$u'(\beta) = \frac{(v(\alpha)+1)B'(\beta)f_h''(h(\beta))}{(v(\alpha)+1)f_h''(h(\beta)) + f_u''(u(\beta))} \tag{EC.5}$$

Under the same assumptions, and additionally by noting that $B'(\beta)$ is positive (Assumption 2), it is clear that both expressions above are positive.

The ratio of FV to OF, $h(\alpha,\beta)/u(\alpha,\beta)$, however, is not necessarily increasing in β. This is demonstrated by considering the derivative of the ratio. The change in the ratio is positive if and only if

$$(v+1)h(\beta)f_h''(h(\beta)) - u(\beta)f_u''(u(\beta)) \geq 0.$$

□

Proof of Corollary 1. Suppose there exists an intervention bundle (α_k,β_k) in the optimal solution to Problem 3 such that $\alpha_k+\beta_k < U$. Let $s \triangleq U-(\alpha_k+\beta_k)$. From Proposition 1, $h_i(\alpha_k+\epsilon_1,\beta_k+\epsilon_2) \geq h_i(\alpha_k,\beta_k)$ for any $\epsilon_1,\epsilon_2 \geq 0$. Therefore, replacing (α_k,β_k) with (α_k',β_k') such that $\alpha_k' = \alpha_k+\delta s$ and $\beta_k' = \beta_k+(1-\delta)s$ for any $\delta \in [0,1]$ maintains feasibility of Problem 3 and satisfies $\alpha_k'+\beta_k' = U$. Additionally, as the total number of intervention bundles does not change as a result of this substitution, this solution is also optimal. □

EC.2.2. Proof of Proposition 2

In the following proof we will drop the subscript i for simplicity. Necessary and sufficient conditions for β and α to be complementary are established by considering the cross derivative $\frac{\partial^2 h(\alpha,\beta)}{\partial\alpha\partial\beta}$. The necessary and sufficient condition for complementary is:

$$\frac{\partial^2 h(\alpha,\beta)}{\partial\alpha\partial\beta} \geq 0.$$

Using the implicit derivatives shown in the proof of Lemma 1, the cross derivative is given by

$$\frac{B'(\beta)\left(f_h''(h)\left((v+1)f_u^{(3)}(u)f_h'(h)-f_u''(u)^2\right)+(v+1)f_h^{(3)}(h)f_h'(h(\alpha,\beta))f_u''(u)-(v+1)f_h''(h)^2f_u''(u)\right)}{(v+1)f_h''(h)+f_u''(u)}. \tag{EC.6}$$

Our goal is to simplify this expression in order to deduce its sign. Notice that, because the denominator is negative, the sign of Expression EC.6 depends entirely on the sign of

$$-f_h''(h)\left((v+1)f_u^{(3)}(u))f_h'(h)-f_u''(u)^2\right)-(v+1)f_h^{(3)}(h)f_h'(h)f_u''(u)+(v+1)f_h''(h)^2f_u''(u).$$

In particular, if this expression is positive then α and β are complementary. Further simplification and manipulation, along with noting that $(1+v)f_h'(h)+f_u'(u)=0$ (by the KKT conditions), yields the necessary and sufficient condition in the proposition.

Now consider specific utility functions. First, consider iso-elastic utility functions. When both $f_h(\cdot)$ and $f_u(\cdot)$ are of this form, the condition of Proposition 2 is always satisfied with a strict inequality. Therefore, education and monetary interventions are strictly complementary.

When $f_h(\cdot)$ and $f_u(\cdot)$ are within the class of exponential utility functions, the inequality of Proposition 2 holds with equality. Therefore, there is no interaction between the education and monetary interventions. □

EC.2.3. Proof of Theorem 1

For simplicity we will drop the subscript i throughout this proof. We will consider the implicit derivatives $h'(\hat{\beta})$ and $u'(\hat{\beta})$, where $\hat{\beta} \triangleq U-\hat{\alpha}$. Notice that showing that $h'(\hat{\alpha})$ is unimodal in $\hat{\alpha}$ is equivalent to showing that $h'(\hat{\beta})$ is unimodal in $\hat{\beta}$. These derivatives are given below.

$$h'(\hat{\beta}) = -\frac{v'(\alpha)f_h'(h(\hat{\beta}))+B'(\hat{\beta})f_u''(u(\hat{\beta}))}{(v(\alpha)+1)f_h''(h(\hat{\beta}))+f_u''(u(\hat{\beta}))} \tag{EC.7a}$$

$$u'(\hat{\beta}) = \frac{-v'(\alpha)f_h'(h(\hat{\beta}))+(v(\alpha)+1)B'(\hat{\beta})f_h''(h(\hat{\beta}))}{(v(\alpha)+1)f_h''(h(\hat{\beta}))+f_u''(u(\hat{\beta}))} \tag{EC.7b}$$

Notice that $u'(\hat{\beta})$ is always positive. This implies that as the planner allocates more money towards the consumer's food budget and less towards education, OF spending consistently increases. In order to prove that $h(\hat{\beta})$ is unimodal in $\hat{\beta}$, we will show that if $h'(\hat{\beta}_1) < 0$ for some $\hat{\beta}_1$, then $h'(\hat{\beta}) < 0$ for all $\hat{\beta} > \hat{\beta}_1$. This implies that once the function $h(\hat{\beta})$ begins to decrease, it will remain decreasing, implying that there can be at most one local maximum.

Notice that the denominator of $h'(\hat{\beta})$ is always negative. Therefore, the sign of the derivative is determined by the numerator. Suppose that $h'(\hat{\beta}_1) < 0$. Consider the derivative of the numerator of Equation EC.7a with respect to $\hat{\beta}$, evaluated at $\hat{\beta}_1$:

$$h'(\hat{\beta}_1)v'(U-\hat{\beta}_1)f_h''(h(\beta_1)) - v''(U-\hat{\beta}_1)f_h'(h(\hat{\beta}_1)) + B'(\hat{\beta}_1)f_u^{(3)}(u(\hat{\beta}_1))u'(\hat{\beta}_1) + B''(\hat{\beta}_1)f_u''(u(\hat{\beta}_1))$$

The quantity above is positive under Assumptions 1 and 2. Note that this holds for any β_1 such that $h'(\hat{\beta}) \leq 0$. Therefore, for any $\hat{\beta} > \hat{\beta}_1$, it holds that the numerator of $h'(\hat{\beta})$ is positive, implying that $h'(\hat{\beta})$ is negative. Thus, $h(\hat{\beta})$ can have at most one local maximum, and can have no local minimum. □

EC.2.4. Proof of Proposition 3

For simplicity we will drop the subscript i throughout this proof. From Expression EC.7a, the optimal value of $\hat{\alpha}$ which maximizes $h(\hat{\alpha})$ must satisfy $v'(\hat{\alpha})f_h'(h(\hat{\alpha})) + B'(U-\hat{\alpha})f_u''(u(\hat{\alpha})) = 0$. By the KKT conditions, we also know that u and h satisfy $(1+v(\hat{\alpha}))f_h'(h(\hat{\alpha})) - f_u'(u(\hat{\alpha})) = 0$.

As defined in the main text, let $A_{f_u}(u) \triangleq \frac{-f_u''(u(\hat{\alpha}))}{f_u'(u(\hat{\alpha}))}$. We can re-write the first expression above as $v'(\hat{\alpha})f_h'(h(\hat{\alpha})) - B'(U-\hat{\alpha})A_{f_u}(u)f_u'(u(\hat{\alpha})) = 0$. Using the KKT conditions to substitute $(1+v(\hat{\alpha}))f_h'(h(\hat{\alpha}))$ for $f_u'(u(\hat{\alpha}))$ yields $v'(\hat{\alpha})f_h'(h(\hat{\alpha})) - B'(U-\hat{\alpha})A_{f_u}(u)(1+v(\hat{\alpha}))f_h'(h(\hat{\alpha})) = 0$ which is equivalent to $v'(\hat{\alpha}) - B'(U-\hat{\alpha})A_{f_u}(u)(1+v(\hat{\alpha})) = 0$. Thus, the optimal value of $\hat{\alpha}$ satisfies $\frac{v'(\hat{\alpha})}{1+v(\hat{\alpha})} = B'(U-\hat{\alpha})A_{f_u}(u)$. □

EC.2.5. Proof of Proposition 4, Theorem 2, and Corollary 2

Before proving Proposition 4 and Theorem 2, we will prove a number of lemmas that will be used in these proofs.

LEMMA EC.1. *On the domain* $\{\hat{\alpha} \in [0,U] : h_i'(\hat{\alpha}) \leq 0\}$,

$$h_i''(\hat{\alpha}) \geq -B_i'(U-\hat{\alpha}^*)\frac{A_{f_{u_i}}(u_i^*)A_{v_i}^+}{A_{f_{h_i}}^-} - B_i'(U-\hat{\alpha})(h_i'(\hat{\alpha}) + B_i'(U-\hat{\alpha}))A_{f_{u_i}'}^+ + B''(U-\hat{\alpha})$$

where $A_f(y) \triangleq \frac{-f''(y)}{f'(y)}$, $A_f^+ \triangleq \max_y \frac{-f''(y)}{f'(y)}$, *and* $A_f^- \triangleq \min_y \frac{-f''(y)}{f'(y)}$.

Proof of Lemma EC.1 For simplicity, throughout the proof we will drop the subscript i. We can calculate $h''(\hat{\alpha})$ through implicit differentiation, as in the proof of Lemma 1. This results in

$$h''(\hat{\alpha}) = \frac{h'(\hat{\alpha})\mathcal{F}(h,u) + v''(\hat{\alpha})f_h'(h(\hat{\alpha})) + f_u^{(3)}(u(\hat{\alpha}))u'(\hat{\alpha})B'(U-\hat{\alpha}) - f_u''(u(\hat{\alpha}))B''(U-\hat{\alpha})}{-f_u''(u(\hat{\alpha})) - (v(\hat{\alpha})+1)f_h''(h(\hat{\alpha}))}$$

where $\mathcal{F}(h,u) = 2v'(\hat{\alpha})f_h''(h(\hat{\alpha})) + (v(\hat{\alpha})+1)f_h^{(3)}(h(\hat{\alpha}))h'(\hat{\alpha}) + f_u^{(3)}(u(\hat{\alpha}))u'(\hat{\alpha})$. Because the denominator is positive, we can construct a lower bound of $h''(\hat{\alpha})$ by keeping only the negative terms in the numerator. In particular, notice that because of Assumption 1 the function $\mathcal{F}(h,u)$ is always negative on the domain $\{\hat{\alpha} \in [0,U]: \; h_i'(\hat{\alpha}) \le 0\}$. Therefore, $h'(\hat{\alpha})\mathcal{F}(h,u)$ is positive, and thus removing this term from the numerator results in the following valid lower bound of $h''(\hat{\alpha})$. We therefore can write the following string of inequalities:

$$\begin{aligned}
h''(\hat{\alpha}) \ge& \frac{f_h(h(\hat{\alpha}))v''(\hat{\alpha}) + B'(U-\hat{\alpha})u'(\hat{\alpha})f_u^{(3)}(u(\hat{\alpha})) - f_u''(u(\hat{\alpha}))B''(U-\hat{\alpha})}{-f_u''(u(\hat{\alpha})) - (v(\hat{\alpha})+1)f_h''(h(\hat{\alpha}))} \\
\ge& \frac{f_h(h(\hat{\alpha}))v''(\hat{\alpha})}{-(v(\hat{\alpha})+1)f_h''(h(\hat{\alpha}))} \\
&- \frac{B'(U-\hat{\alpha})(h'(\hat{\alpha}) + B'(U-\hat{\alpha}))f_u^{(3)}(u(\hat{\alpha})) + f_u''(u(\hat{\alpha}))B''(U-\hat{\alpha})}{-f_u''(u(\hat{\alpha}))} \\
\ge& \frac{v''(\hat{\alpha})}{v(\hat{\alpha})+1} \max_{h(\hat{\alpha}):\,\hat{\alpha}\in[0,U]} \frac{f_h(h)}{|f_h''(h)|} + B''(U-\hat{\alpha}) \\
&- B'(U-\hat{\alpha})(h'(\hat{\alpha}) + B'(U-\hat{\alpha})) \max_{u(\hat{\alpha}):\,\hat{\alpha}\in[0,U]} \frac{f_u^{(3)}(u)}{|f_u''(u)|} \\
\ge& \frac{v''(\hat{\alpha})}{v(\hat{\alpha})+1} \frac{1}{\min_{h(\hat{\alpha}):\,\hat{\alpha}\in[0,U]} A_{f_h}(h)} + B''(U-\hat{\alpha}) \\
&- B'(U-\hat{\alpha})(h'(\hat{\alpha}) + B'(U-\hat{\alpha})) \max_{u(\hat{\alpha}):\,\hat{\alpha}\in[0,U]} A_{f_u'}(u) \\
=& \frac{v''(\hat{\alpha})}{v(\hat{\alpha})+1} \frac{1}{A_{f_h}^-} - B'(U-\hat{\alpha})(h'(\hat{\alpha}) + B'(U-\hat{\alpha}))A_{f_u'}^+ + B''(U-\hat{\alpha}) \\
=& \frac{v'(\hat{\alpha})}{v(\hat{\alpha})+1} \frac{v''(\hat{\alpha})}{v'(\hat{\alpha})} \frac{1}{A_{f_h}^-} - B'(U-\hat{\alpha})(h'(\hat{\alpha}) + B'(U-\hat{\alpha}))A_{f_u'}^+ + B''(U-\hat{\alpha})
\end{aligned}$$

The second inequality follows by separating the expression into two terms and removing positive terms from the denominators. The third inequality follows by taking the maximum of the ratios, and the fourth inequality and following equality follow by substituting in the appropriate notation.

Simplifying further, we can write

$$\begin{aligned} h''(\hat\alpha) &\geq -\frac{v'(\hat\alpha)}{v(\hat\alpha)+1}\frac{A_v^+}{A_{f_h}^-} - B'(U-\hat\alpha)(h'(\hat\alpha)+B'(U-\hat\alpha))A^+_{f'_u} + B''(U-\hat\alpha) \\ &\geq -\frac{v'(\hat\alpha^*)}{v(\hat\alpha^*)+1}\frac{A_v^+}{A_{f_h}^-} - B'(U-\hat\alpha)(h'(\hat\alpha)+B'(U-\hat\alpha))A^+_{f'_u} + B''(U-\hat\alpha) \\ &= -B'(U-\hat\alpha^*)\frac{A_{f_u}(u_i^*)A_v^+}{A_{f_h}^-} - B'(U-\hat\alpha)(h'(\hat\alpha)+B'(U-\hat\alpha))A^+_{f'_u} + B''(U-\hat\alpha) \end{aligned}$$

The first inequality lower bounds $\frac{v''(\hat\alpha)}{v'(\alpha)}$ by $-A_v^+$. The second inequality follows by the concavity of $v(\hat\alpha)$ and the fact that we are focused on the domain where $\hat\alpha \geq \hat\alpha^*$, and the final equality follows by Proposition 3. □

Lemma EC.2. *The function*

$$y(\hat\alpha) \triangleq h^* + \frac{c_2}{c_1^2}\left(1 + c_1(\hat\alpha^* - \hat\alpha) - e^{c_1(\alpha^* - \hat\alpha)}\right)$$

is lower bounded by

$$\underline{y}(\alpha) \triangleq h^* - \frac{c_2}{2}(\hat\alpha - \hat\alpha^*)^2$$

on the domain $[\hat\alpha^*, 1]$.

Proof of Lemma EC.2 Consider the second-order Taylor expansion of $y(\hat\alpha)$, given by

$$h^* - \frac{1}{2}c_2(\hat\alpha - \hat\alpha^*)^2 + O\left((\hat\alpha - \hat\alpha^*)^3\right).$$

The error is exactly given by $\sum_{k=3}^{\infty} t^k(\hat\alpha)$ where $t^k(\hat\alpha)$ is the t-th order Taylor term. In this case, it is given by $c_2\sum_{k=3}^{\infty}\frac{(-1)^{k-1}(x-\hat\alpha^*)^k c_1^{k-2}}{k!} = c_2\left(\frac{2-2e^{-c_1(x-\hat\alpha^*)}}{c_1^2} - \frac{x-\hat\alpha^*}{c_1} + \frac{1}{2}(x-\hat\alpha^*)^2\right)$. The error term is increasing in $x - \hat\alpha^*$ and of course equal to zero when $x = \hat\alpha^*$. Therefore, on $[0, \hat\alpha^*]$ the error is negative and on $[\hat\alpha^*, 1]$ the error is positive. Thus, on $[\hat\alpha^*, 1]$ the second-order approximation is already a valid lower bound since the error is positive. □

We are now ready to prove Proposition 4. To simplify the exposition we will use the notation $A_f(y) \triangleq \frac{-f''(y)}{f'(y)}$, $A_f^+ \triangleq \max_y \frac{-f''(y)}{f'(y)}$, and $A_f^- \triangleq \min_y \frac{-f''(y)}{f'(y)}$.

Proof of Proposition 4. First, we will focus on constructing the lower bound, $\underline{h}_i(\hat\alpha)$, of $h_i(\hat\alpha)$ for an individual i. Our goal is to construct a tractable lower bound that is "close" to $h(\hat\alpha)$ at its peak. We will construct two lower bounds, one that is valid on the domain $\mathcal{D}_1 \triangleq \{\hat\alpha :\ h'(\hat\alpha) \geq 0\}$ and one that is valid on $\mathcal{D}_2 \triangleq \{\hat\alpha :\ h'(\hat\alpha) \leq 0\}$. Note that by Theorem 1 above, these domains are continuous.

The lower bound on $\mathcal{D}_1$ is trivial: we set the lower bound to be identically zero. Thus, we focus on the lower bound on $\mathcal{D}_2$. On domain $\mathcal{D}_2$ we will begin by constructing a lower bound of $h''(\hat{\alpha})$, and then use this lower bound to solve a second-order differential equation to obtain a lower bound of $h(\hat{\alpha})$. As $h''(\hat{\alpha})$ is negative on $\mathcal{D}_2$, the lower bound will be a concave decreasing function that matches the function $h(\hat{\alpha})$ at α^*.

Lemma EC.1 shows that on $\mathcal{D}_2$,

$$h_i''(\hat{\alpha}) \geq -B'(U-\hat{\alpha}^*)\frac{A_{f_u}(u_i^*)A_v^+}{A_{f_h}^-} - B'(U-\hat{\alpha})(h_i'(\hat{\alpha})+B'(U-\hat{\alpha}))A_{f_u'}^+ + B''(U-\hat{\alpha}).$$

Let $\hat{\beta} := U - \hat{\alpha}$. Recalling that $B(\hat{\beta})$ is concave increasing, and that $\hat{\alpha} \geq \hat{\alpha}^*$ on $\mathcal{D}_1$, we know that $B'(\hat{\beta}) \geq B'(\hat{\beta}^*)$ since $\hat{\beta} \leq \hat{\beta}^*$ on $\mathcal{D}_1$. From Equation EC.7a, we know that

$$h'(\hat{\alpha}) = -\frac{v'(\hat{\alpha})f_h'(h(\hat{\beta})) + B'(\hat{\beta})f_u''(u(\hat{\beta}))}{(v+1)f_h''(h(\hat{\beta})) + f_u''(u(\hat{\beta}))}.$$

Therefore, $h'(\hat{\alpha}) \geq -\frac{B'(\hat{\beta})f_u''(u(\hat{\beta}))}{(v+1)f_h''(h(\hat{\beta}))+f_u''(u(\hat{\beta}))} \geq -B'(\hat{\beta})$ and thus $h'(\hat{\alpha}) + B'(\hat{\beta}) \geq 0$. Therefore, since each term in the lower bound of $h_i''(\hat{\alpha})$ is negative, we can obtain another valid lower bound by replacing $B'(U-\hat{\alpha}^*)$ and $B'(U-\hat{\alpha})$ with their largest possible values, and replacing $B''(U-\hat{\alpha})$ with its smallest possible value. The new lower bound therefore becomes

$$h_i''(\hat{\alpha}) \geq -B'(0)\frac{A_{f_u}(u_i^*)A_v^+}{A_{f_h}^-} - B'(0)(h_i'(\hat{\alpha})+B'(0))A_{f_u'}^+ + \min_{\beta\in[0,U]} B''(\beta).$$

Let the right-hand side of the above inequality be denoted by the function $L(\hat{\alpha}, h_i')$.

This lower bound can be used to construct a lower bound of $h(\hat{\alpha})$ by solving the following differential equation with initial conditions $\underline{h}'(\hat{\alpha}^*) = 0$ and $\underline{h}(\hat{\alpha}^*) = h^*$:

$$\begin{cases} \underline{h}''(\hat{\alpha}) = L(\hat{\alpha}, \underline{h}'(\hat{\alpha})) \\ \underline{h}'(\hat{\alpha}^*) = 0 \\ \underline{h}(\hat{\alpha}^*) = h^*. \end{cases} \tag{EC.8}$$

The solution to the differential equation is:

$$\tilde{\underline{h}} = h^* + \frac{l(\hat{\alpha}_i^*)}{d^2}\left(1 + d\cdot(\hat{\alpha}^* - \hat{\alpha}) - e^{d\hat{\alpha}(\alpha^*-\hat{\alpha})}\right)$$

where $d \triangleq A^+_{f'_u} B'(0)$ and $l(\hat{\alpha}^*_i) \triangleq B'(0)\frac{A_{f_u}(u^*_i)A^+_v}{A^-_{f_h}} + B'(0)^2 A^+_{f'_u} - \min_{\beta\in[0,U]} B''(\beta)$. Notice that this definition of $l(\hat{\alpha}^*_i)$ is equivalent to that of Proposition 4.

By Lemma EC.2, the function

$$\underline{h}(\alpha) \triangleq h(\hat{\alpha}^*) - \frac{l(\hat{\alpha}^*_i)}{2}(\hat{\alpha} - \hat{\alpha}^*)^2 \tag{EC.9}$$

is a valid lower bound of $\underline{\tilde{h}}(\alpha)$. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. We will begin by proving part (1). First, we remark that because of Assumption 4, q_i and Q_i exist for all $i \in \mathcal{P}$. We will show that for any κ such that $\kappa \le (Q_i - q_i)$ for all $i \in \mathcal{S}$, $\mathcal{M}(\kappa)$ constructed as in part (1) ensures that there exists an intervention $\hat{\alpha}_{k(i)} \in \mathcal{M}(\kappa)$ such that $h_i(\hat{\alpha}_{k(i)}) \ge \tau_i$ for all $i \in \mathcal{P}$. First note that $\mathcal{M}(\kappa)$ contains one intervention bundle in every length-κ interval of the domain $[0,U]$, including an intervention bundle at the endpoints $\hat{\alpha} = 0$ and $\hat{\alpha} = U$. This results in exactly $1 + \lceil U/\kappa \rceil$ intervention bundles.

Consider an individual $i \in \mathcal{P}$. For this person, let $\mathcal{L}_i \triangleq \{\hat{\alpha} \in [0,U] : h_i(\hat{\alpha}) \ge \tau_i\}$ be the set of intervention bundles such that their FV spending exceeds τ_i, and let $\bar{q}_i \triangleq Q_i - q_i = |\mathcal{L}_i|$ be the length of the closed interval $\mathcal{L}_i$ (which is positive by Assumption 4). For person i, there exists an intervention $\hat{\alpha}_{k(i)} \in \mathcal{M}(\kappa)$ such that $h_i(\hat{\alpha}_{k(i)}) \ge \tau_i$ if and only if $\mathcal{L}_i \cap \mathcal{M}(\kappa) \ne \emptyset$. Notice that we use the notation $\hat{\alpha}_{k(i)} \in \mathcal{M}(\kappa)$ as shorthand for $(\hat{\alpha}_{k(i)}, U - \hat{\alpha}_{k(i)}) \in \mathcal{M}(\kappa)$

Suppose $i \notin \mathcal{S}$. This implies that either $0 \in \mathcal{L}_i$ or $U \in \mathcal{L}_i$. Since $0, U \in \mathcal{M}(\kappa)$ by construction, individual i can either be assigned to $\hat{\alpha} = 0$ or $\hat{\alpha} = U$ and achieve FV spending above τ_i.

Now suppose that $i \in \mathcal{S}$. By definition, $\kappa \le \bar{q}_i$ and there exists an intervention within every length-κ interval in $[0,U]$. Therefore, $\mathcal{L}_i$ must contain one such intervention bundle.

Now we will prove part (2). We will use the lower bounds constructed in the proof of Proposition 4 to characterize the level set $\{\hat{\alpha} : h_i(\hat{\alpha}) \ge \tau_i\}$. In particular, we will find a subset of this level set by computing values $z_i^{(1)}$ and $z_i^{(2)}$ that $\{\hat{\alpha} \in [z_i^{(1)}, z_i^{(2)}]\} \subset \mathcal{L}_i$. We will set $z_i^{(1)} = \hat{\alpha}^*_i$ and compute $z_i^{(2)}$ by solving the equation $\underline{h}_i(\hat{\alpha}) = \tau_i$ for i such that $U \notin \mathcal{L}_i$. This results in $z_i^{(2)} = \hat{\alpha}^*_i + \frac{\sqrt{2}\sqrt{h_i(\hat{\alpha}^*_i) - \tau_i}}{\sqrt{l(\hat{\alpha}^*_i)}}$.

The length of the segment $[z_{1_i}, z_{2_i}]$, denoted by $w_i \triangleq z_{2_i} - z_{1_i}$, is therefore a lower bound of $\bar{q}_i$, given by

$$w_i = \frac{\sqrt{2(h_i(\hat{\alpha}^*_i) - \tau_i)}}{\sqrt{l(\hat{\alpha}^*_i)}} \tag{EC.10}$$

where $l(\hat{\alpha}_i^*)$ is defined in the proof of Proposition 4 as $l(\hat{\alpha}_i^*) \triangleq B'(0)\frac{A_{f_u}(u_i^*)A_v^+}{A_{f_h}^-} + B'(0)^2 A_{f_u'}^+ - \min_{\beta\in[0,U]} B''(\beta)$.

Thus, as $w_i \le \bar{q}_i$ for all $i \in \mathcal{S}$, setting $\kappa = w^{\min} \triangleq \min_{i\in\mathcal{S}} w_i$ is a valid choice.

□

Proof of Corollary 2. For any individual i, the proof of Theorem 2 shows that any $\hat{\alpha} \in [\hat{\alpha}_i^*, \hat{\alpha}_i^* + w^{\min}]$ will result in FV spending at least as large as τ_i. By the definition of $\mathcal{M}(\kappa)$ with $\kappa = w^{\min}$, there is guaranteed to be at least one intervention in every interval of length $w^{\min}$. Therefore, assigning consumer i to $\hat{\alpha}_{k(i)}$ where $\hat{\alpha}_{k(i)} = \arg\min\{\hat{\alpha} \in \mathcal{M}(\kappa) : \hat{\alpha} \ge \hat{\alpha}_i^*\}$ achieves the desired level of FV spending. □

EC.2.6. Proof of Example 1.

When $f_{u_i}(\cdot)$ and $f_{h_i}(\cdot)$ are exponential we can write

$$h_i(\hat{\alpha}) = \frac{\log\left(\frac{\eta_{h_i}(v_i(\hat{\alpha})+1)}{\eta_{u_i} c_{u_i}}\right) + \eta_{u_i} B_i(U-\hat{\alpha})}{\eta_{h_i} + \eta_{u_i}},$$

and

$$h_i'(\hat{\alpha}) = \frac{-\eta_{u_i} B_i'(U-\hat{\alpha}) + \frac{v_i'(\hat{\alpha})}{1+v_{i_0}}}{\eta_{u_i} + \eta_{h_i}}.$$

Let $B_i'(\beta) = b_i$ as in the example. First consider person j, for which $v_{i_0} = 1$. For this person, $v_i'(\hat{\alpha}) = 0$ and therefore $h_j'(\hat{\alpha}) = -\eta_{u_j} b_j/(\eta_{u_j} + \eta_{h_j})$. For person j, the best intervention bundle is $\hat{\alpha}_j^* = 0$ which results in $h_j(\hat{\alpha}_j^*) = h_{j_0} + \frac{\eta_{u_i} U b_j}{\eta_{u_j}+\eta_{h_j}}$. In general, for an intervention bundle $\hat{\alpha}$, $h_j(\hat{\alpha}) = h_{j_0} + \frac{\eta_{u_i}\hat{\alpha} b_j}{\eta_{u_j}+\eta_{h_j}}$ because $v_j'(\hat{\alpha}) = 0$. Therefore, $h_j(\hat{\alpha}) \ge h_{i_0} + \rho(h_i(\hat{\alpha}_i^*) - h_{i_0})$ if and only if $\hat{\alpha} \le U(1-\rho)$.

Now consider person k. This person is impacted by both the monetary and education intervention, since $v_{k_0} = 0$. For this person, $h_k(\hat{\alpha}) = \frac{\log\left(\frac{\eta_{h_k}(\nu_k\hat{\alpha}+1)}{\eta_{u_k} c_{u_k}}\right) + \eta_{u_k} B_k(U-\hat{\alpha})}{\eta_{h_k}+\eta_{u_k}}$. Since $\nu_k > b_k \eta_{u_k}$, $h_i'(\hat{\alpha})$ is positive and therefore $\hat{\alpha}_k^* = U$. The inequality $h_k(\hat{\alpha}) \ge h_{k_0} + \rho(h_k(\hat{\alpha}_k^*) - h_{k_0})$ can be written as:

$$\eta_{u_k} b_k(U-\hat{\alpha}) + \log(1+\hat{\alpha}\nu_k) \ge \rho\log(1+\nu_k U) \tag{EC.11}$$

We wish to find a lower bound on the minimum value of $\hat{\alpha}$ such that the above inequality holds. Consider a second inequality given by:

$$\eta_{u_k} b_k(U-\hat{\alpha}) + \hat{\alpha}\nu_k \ge \rho\log(1+\nu_k U) \tag{EC.12}$$

Note that for all $x \geq -1$, the inequality $\log(1+x) \leq x$ holds. Therefore, the smallest value of $\hat{\alpha}$ such that inequality EC.12 holds must be smaller than the smallest $\hat{\alpha}$ such that inequality EC.11 holds, so solving EC.12 results in a lower bound on the feasible $\hat{\alpha}$ for EC.11. The solution to EC.12 is

$$\hat{\alpha} \geq \frac{\rho \log(1+\nu_k U) - b_k U \eta_{u_k}}{\nu_k - \eta_{u_k} b_k}.$$

Putting these pieces together,we wish to understand when there exists a choice of $\hat{\alpha}$ such that *both* $\hat{\alpha} \leq U(1-\rho)$ and $\hat{\alpha} \geq \frac{\rho \log(1+\nu_k U) - bKU\eta_{u_k}}{\nu_k - \eta_{u_k} b_k}$ are true at the same time. When $U(1-\rho) < \frac{\rho \log(1+\nu_k U) - bKU\eta_{u_k}}{\nu_k - \eta_{u_k} b_k}$, no such $\hat{\alpha}$ exists. This occurs when $\rho \geq \frac{U\eta_k}{U(\nu_k - b_k \eta_{u_k}) + \log(1+U\nu_k)}$, which is the inequality stated in the proposition.

Thus, when $\rho \geq \frac{U\eta_k}{U(\nu_k - b_k \eta_{u_k}) + \log(1+U\nu_k)}$, there is no single intervention bundle $\hat{\alpha}$ that can ensure both $h_j(\hat{\alpha}) \geq \tau_j$ and $h_k(\hat{\alpha}) \geq \tau_k$ simultaneously, where τ_i is defined as in the statement of the proposition.

□

EC.2.7. Proposition EC.1

PROPOSITION EC.1. *Under Assumption EC.1, stated below, the following holds: When utility functions are exponential (isoelastic), combining the monetary and education interventions can increase the FV spending achieved by either individual intervention by 75% (180%). The percentage increase is defined as* $\frac{h(\hat{\alpha}_i^*) - \max\{h_i(0), h_i(U)\}}{\max\{h_i(0), h_i(U)\}}$.

ASSUMPTION EC.1. *The following conditions narrow down the parameter domain:*

- *The value of nutrition linking function satisfies* $v_{0_i} + 0.1(1 - v_{0_i}) \geq v_i(1/2)$ *and* $v_i(1/2) + 0.1(1 - v_i(1/2)) \geq v_i(1)$ *for all* i
- *The monetary linking function satisfies* $B_i(1/2) \geq B_{0_i} + 1$, $B_i(1) \geq B_i(1/2) + 1$ *and* $B_i(1) \leq B_{i_0} + 10$ *for all* i
- *When utility functions are exponential,* $\eta_{u_i}, \eta_{h_i} \in [0.005, 0.06]$, $c_u \in [.1, 5]$, $c_h = 1$, *and* $\eta_{u_i} \geq \eta_{h_i}$.
- *When utility functions are isoelastic,* $\eta_{u_i}, \eta_{h_i} \in [0.1, 0.95]$, $c_u \in [.1, 5]$, $c_h = 1$, *and* $\eta_{u_i} \geq \eta_{h_i}$.

We note that setting $c_h = 1$ *is without loss of generality.*

Assumption EC.1 defines a reasonable parameter space in which both linking functions make intuitive sense, as do the utility functions.The condition that $\eta_{u_i} \geq \eta_{h_i}$ ensures that the utility function for FVs is more concave than the utility function for OFs, as should be the case in reality.

Proof of Proposition EC.1. We prove this result by giving a particular set of parameters that achieves the desired lower bound. In the case of exponential utility, the parameters are given in Table EC.2 and for isoelastic utility functions they are given in Table EC.3.

η_h	η_u	c_u	B_0	$B(1/2)$	$B(1)$	v_0	$v(1/2)$	$v(1)$
0.036	0.036	4.304	33.038	42.038	43.038	0.389	1.0	1.0

Table EC.2 Parameters that achieve the desired lower bound when utility functions are exponential.

η_h	η_u	c_u	B_0	$B(1/2)$	$B(1)$	v_0	$v(1/2)$	$v(1)$
0.138	0.103	1.621	3.611	12.611	13.611	0.561	1.0	1.0

Table EC.3 Parameters that achieve the desired lower bound when utility functions are isoelastic.

□

EC.3. Alternative planner objectives

In this section we discuss extensions of our model and main results to consider alternative objective functions for the planner. In the main text, we assume that the planner's goal is to minimize the number of unique intervention bundles created, corresponding to fixed costs. In this section, we explore the tradeoff between fixed costs and variable costs, in addition to considering different fixed costs for each intervention. The variable cost of an intervention (α, β) is $\alpha + \beta$, whereas the fixed cost is proportional to the number of unique intervention bundles utilized.

Our goal is to extend Theorem 2 to a more general settings where variable cost is taken into consideration. In the main text, every intervention bundle is restricted so that $\alpha + \beta \leq U$, and the planner does consider variable costs in their optimization problem. Thus, every bundle will satisfy $\alpha + \beta = U$. At a high-level, the main difficulty for the planner in minimizing variable costs is that meeting the condition $h_i(\alpha, \beta) \geq \tau_i$ becomes more difficult as the per-person budget becomes more constrained. Figure EC.1 illustrates this idea. When the budget is equal to U, the thick blue line in Figure EC.1 shows the range of intervention bundles that would achieve $h_i(\alpha, \beta) \geq \tau_i$. As the total budget shrinks,

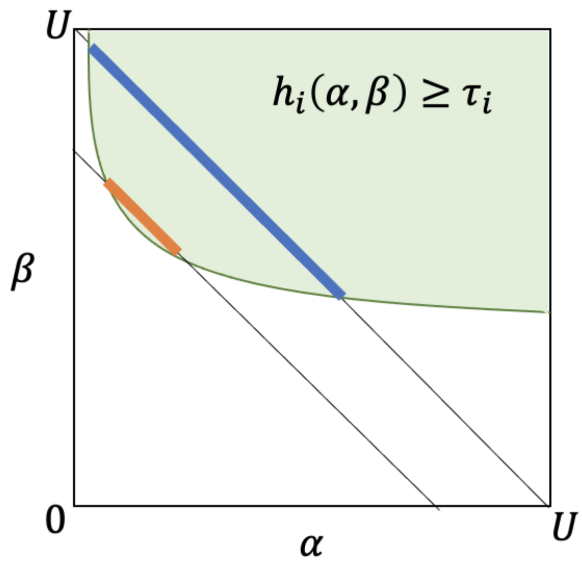

Figure EC.1 **Illustration of the tradeoff between variable cost and sensitivity.**

this range decreases. The thick orange line shows an example of the new range of bundles that will achieve $h_i(\alpha, \beta) \geq \tau_i$ with a smaller budget.

Section EC.3.1 discusses two models that allow variable cost to be a decision of the planner. In the first model, all individuals still receive an equal total investment, but this total amount is a decision of the planner. In the second model, the planner can divide the population into two tiers that each receive different investment amounts. Then, Section EC.3.2 considers a model in which only distinct education intervention levels incur a fixed cost (i.e., the fixed cost of varying the monetary intervention is zero), but both intervention types have variable costs associated with them. Finally, Section EC.3.3 discusses the practicality of implementing these approaches given that individual-level utility functions are unknown in reality.

EC.3.1. Tradeoff between fixed and variable cost

In the main text, each person receives a intervention bundle with total cost U, and thus the total variable cost is UP. If the planner wished to spend less than U dollars on every individual, they would likely need to increase the level of personalization to compensate for the more constrained budget, as illustrated in Figure EC.1. In this section, we consider such a tradeoff by considering the following problem for the planner:

$$\begin{aligned} \min_{\substack{M,\,\{(\alpha_k,\beta_k)\}_{k\in\{1,\dots,M\}} \\ z_{ik}\,\forall\, i\in[P],k\in[K]}} \quad & M+\gamma\sum_{i\in\mathcal{P}}\left(\sum_{k=1}^{K} z_{ik}(\alpha_k+\beta_k)\right) && \text{(Problem EC.1)} \\ s.t.\; & \beta_k+\alpha_k \le U && \forall\, k\in[M] \\ & h_i\left(\sum_{k=1}^{K} z_{ik}\alpha_k, \sum_{k=1}^{K} z_{ik}\beta_k\right) \ge \tau_i && \forall i\in[P] \\ & z_{ik}\in\{0,1\} && \forall\, i\in[P],\, k\in[M] \\ & \sum_{k=1}^{K} z_{ik} = 1 && \forall\, i\in[P] \\ & \beta_k,\, \alpha_k \ge 0 && \forall\, k\in[M] \end{aligned}$$

This differs from Problem 3 in that we now include the total variable cost in the objective function, and use γ to weigh the tradeoff between fixed and variable costs.

When only fixed costs were considered, understanding the necessary number of bundles depended on κ, which needs to be a lower bound of $Q_i - q_i$ for all $i \in \mathcal{S}$ (see Theorem 2). Now, we introduce related notation that will allow us to consider the tradeoff between fixed and variable costs. Specifically, let H_i be the maximal hypotenuse length of an isosceles right triangle inscribed in $\{\alpha, \beta : h_i(\alpha, \beta) \ge \tau_i,\ \alpha + \beta \le U\}$. This is illustrated in Figure EC.2.

EC.3.1.1. Single tier Now, we begin by considering the case where each individual still receives the same investment, but we allow the investment amount to be a decision variable. We assume that every individual receives a total investment amount of $U - \delta$. When $\delta > 0$ the total variable cost is decreased compared to the formulation in the main text. Given the same thresholds τ_i for $i \in \mathcal{P}$, a decrease in the total variable cost may also require an increase the number of intervention bundles in order to maintain an outcome of at least τ_i for individual i. In other words, as the budget becomes more constrained, it is increasingly important that each person receive a customized intervention bundle. In this setting, Theorem EC.1 characterizes the tradeoff between fixed costs, total variable cost, and the thresholds $\boldsymbol{\tau} = \{\tau_i\}_{i\in\mathcal{P}}$.

Theorem EC.1. *For $\delta \in [0, \frac{H^{min}}{\sqrt{2}})$, it is possible to reduce the total variable cost to $(U-\delta)P$ with and use no more than $1 + \lceil \frac{\sqrt{2}(U-\delta)}{H^{min} - \delta\sqrt{2}} \rceil$ intervention bundles, where $H^{min} \triangleq \min_{i\in\mathcal{P}} H_i$.*

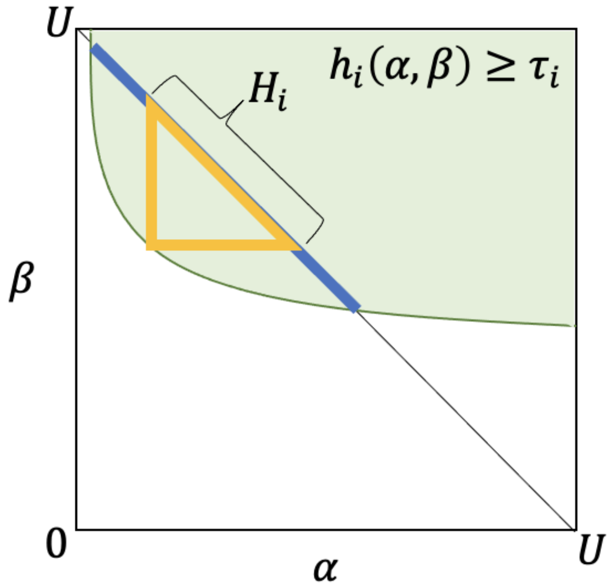

Figure EC.2 **Illustration of H_i.**

The intuition for Theorem EC.1 is the following: As δ increases, we simultaneously reduce the cost of each person's bundle, but also increase the population's sensitivity to their given bundle. The quantity $\frac{H^{min}}{\sqrt{2}}$ is the maximum amount that the planner's can decrease *everyone's* variable cost while ensuring the feasibility of Problem EC.1.

Proof of Theorem EC.1 By the definition of H_i, for person i we can reduce U to up to $U - H_i/\sqrt{2}$ while still having a feasible intervention bundle such that person i's FV spending can be at least τ_i. When U is reduced to $U - \delta$, there exists a continuous segment of length at least $H_i - \delta\sqrt{2}$ on the domain $\{\alpha, \beta : \alpha + \beta = U - \delta, h_i(\alpha, \beta) \geq \tau_i\}$.

Therefore, by dividing the domain $\{\alpha, \beta : \alpha + \beta = U - \delta\}$ into segments of length $H_i - \delta\sqrt{2}$, it guarantees that there will exist at least one intervention such that $h_i(\alpha, \beta) \geq \tau_i$. If we replace H_i with H^{min}, the above argument shows that there is always an intervention for each person in the population.

Lastly, notice that dividing the line segment $\{\alpha, \beta : \alpha + \beta = U - \delta\}$ (which has length $\sqrt{2}(U - \delta)$) into segments of length $H^{min} - \delta\sqrt{2}$ results in $1 + \lceil \frac{\sqrt{2}(U-\delta)}{H^{min}-\delta\sqrt{2}} \rceil$ intervention bundles. □

EC.3.1.2. Two-tiers Next, we consider a second approach whereby the planner constructs two tiers of interventions. Those in Tier 1 receive a total intervention amount of $U - \delta_1$, and those in Tier 2 receive $U - \delta_2$. By strategically segmenting the population into tiers based on H_i, we can put the less sensitive consumers—i.e., those with larger values of H_i—in Tier 2. Specifically, we will choose a threshold $\tilde{H}$. If $H_i \leq \tilde{H}$, individual i is in Tier 1. If $H_i > \tilde{H}$, individual i is in Tier 2. With fixed tiers, we can optimize δ_1 and δ_2. Using these optimal values of δ_1 and δ_2, we can then optimize the threshold $\tilde{H}$. Let $P_2(\tilde{H})$ be the number of people assigned to Tier 2 given a threshold of $\tilde{H}$.

Consider minimizing the following expression:

$$\gamma\Big(P_2(\tilde{H})(U-\delta_2)+(P-P_2(\tilde{H}))(U-\delta_1)\Big)+1+\frac{\sqrt{2}(U-\delta_2)}{\tilde{H}-\sqrt{2}\delta_2}+1+\frac{\sqrt{2}(U-\delta_1)}{H^{min}-\sqrt{2}\delta_1} \quad \text{(EC.13)}$$

The first term is the total variable cost of the proposed scheme, and the last two terms are a continuous approximation of the total fixed cost. These expressions follow from Theorem EC.1. For those in Tier 1, H^{min} remains the lower bound on H_i. For those in Tier 2, $\tilde{H}$ is the lower bound on h_i.

Notice that the Tier 1 and Tier 2 costs are linked through $\tilde{H}$ only. Thus, by fixing $\tilde{H}$, we can compute the optimal values of δ_1 and δ_2 as a function of $\tilde{H}$. These optimal values are given in the following lemma.

LEMMA EC.3. *For a given threshold $\tilde{H}$, the optimal choice for δ_1 and δ_2 that minimize Equation EC.13 is given by*

$$\delta_1=\frac{(P-P_2(\tilde{H}))H^{min}+\sqrt{(2U-\sqrt{2}H^{min})(P-P_2(\tilde{H}))}}{\sqrt{2\gamma}(P-P_2(\tilde{H}))} \quad \text{(EC.14)}$$

$$\delta_2=\frac{\tilde{H}P_2(\tilde{H})-\sqrt{(2U-\sqrt{2}\tilde{H})P_2(\tilde{H})}}{\sqrt{2\gamma}P_2(\tilde{H})} \quad \text{(EC.15)}$$

Using a one- and two-tiered scheme, Figure EC.3 illustrates the tradeoff between fixed and variable cost. In this setting, we assume that utility functions are exponential and use one of the pairs of ground truth utility functions for each individual. Applying Theorem 2 with $U=19$, $\tau_i = h_{i_0} + 0.8(h_i(\hat{\alpha}_i^*) - h_{i_0})$, $\kappa = l^{\min}$, and the linking functions specified in Equations 5 and 6 with $b=0.26$ and $\eta_v=5$ results in four intervention bundles. With one tier, we can reduce the variable cost by about 1\$ per person and still use only four intervention bundles. We could also reduce the per-person cost to about \$15 per person and use 10 intervention bundles. To achieve variable costs below \$15, it is best to use two tiers. For the two tier results shown in Figure EC.3, we compute the optimal threshold $\tilde{H}$ by assuming that the distribution of $H_i - H^{min}$ is exponential with $\lambda = 0.219$. This is a close approximation of the true empirical distribution.

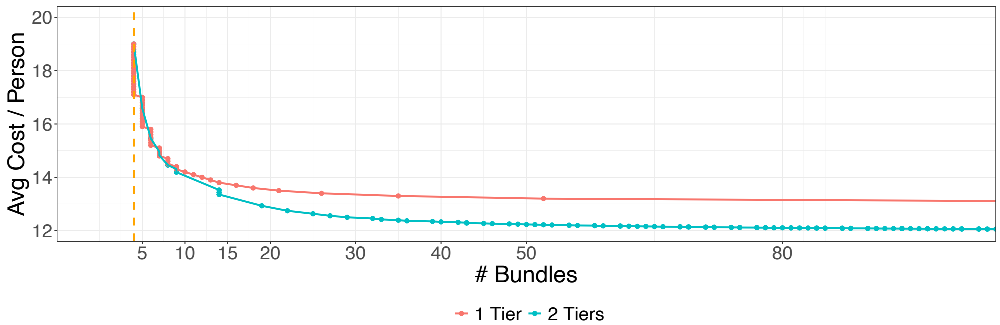

Figure EC.3 Trade-off between fixed and variable cost using one tier and Theorem EC.1 and the two tier method discussed above. The dashed orange line shows the number of bundles that is obtained when $\delta = 0$, using the empirical value of l^{min} in Theorem 2.

EC.3.2. Fixed cost for education interventions only

In this section we consider a variant of the planner's problem in which developing distinct education programs incurs a fixed cost, but varying the values of the monetary intervention does not incur any fixed cost. Both interventions also incur a variable cost. This setting may apply to certain real-world scenarios in which the development of education programs with different levels of investment and effectiveness is highly burdensome, but deploying different monetary intervention levels is not. Varying benefit amounts may, however, decrease the explainability of the program, and thus having too many distinct intervention bundles may still be undesirable from a social perspective.

In this setting, if individual i receives an investment of β_i dollars in the monetary intervention and α_i dollars in the education intervention, $U_i = \alpha_i + \beta_i$ is the planner's total spending on individual i. As before, the quantity $\sum_{i\in\mathcal{P}} U_i$ is the planner's total variable cost. We will let M_α denote the number of distinct education programs offered. The planner wishes to minimize a weighted sum of M_α and $\sum_{i=1}^{P} U_i$.

Our proposed solution for this setting builds directly on Theorem 2. Suppose that Theorem 2 results in a set of intervention bundles $\mathcal{M}$ with cardinality M. In this section we discuss a new solution that results in M unique education intervention levels, but a lower variable cost than the solution of Theorem 2. This is accomplished by varying the investment in the monetary intervention per person.

Let $\hat{\alpha}_i^*(U_i)$ be the optimal intervention bundle for individual i when the total investment for individual i is U_i, and let $\beta_i^*(U_i) = U_i - \hat{\alpha}_i^*(U_i)$ be the resulting investment in the monetary intervention. Thus, $\hat{\alpha}_i^*(U)$ is equivalent to $\hat{\alpha}_i^*$ in the main text. Additionally, let $h_i^*(U_i)$ be individual i's optimal FV spending when the total investment for individual i is

U_i, so $h_i^*(U)$ is equivalent to $h_i(\hat{\alpha}_i^*)$ in the main text. The following corollary is critical to the proposed solution method, and follows from Proposition 3.

COROLLARY EC.1. *When $B_i(\beta)$ is linear and $f_{u_i}(u)$ is exponential, $\frac{\partial \hat{\alpha}_i^*(U_i)}{\partial U_i} = 0$ on the domain $\{U_i : \hat{\alpha}_i^*(U_i) \in [0, U_i)\}$.*

By Corollary EC.1, if $\hat{\alpha}_i^*(U_i)$ is optimal for a certain value of U_i, then, for all $U_i' \neq U_i$, $\hat{\alpha}_i^*(U_i') = \hat{\alpha}_i^*(U_i)$ as long as $\hat{\alpha}_i^*(U_i) < U_i$ and $\hat{\alpha}_i^*(U_i') < U_i'$. In words, Corollary EC.1 says that the optimal investment in education for individual i is fixed regardless of the planner's total investment for individual i.

This suggests that a reasonable strategy for minimizing the investment for individual i might be to fix the investment in the monetary intervention to $\hat{\alpha}_i^*(U)$ and then decrease U_i (thereby decreasing β_i^*) to the extent possible. The following theorem specifies the amount by which U_i could be reduced while still ensuring that FV spending exceeds τ_i.

THEOREM EC.2. *Suppose that $f_{u_i}(u)$ is exponential. Let $\mathcal{M}(\kappa)$ be constructed according to Theorem 2 with $\kappa = w^{min}$. Then as long as $U_i \geq \max\{\hat{\alpha}_i^*, U - (h_i^*(U) - \tau_i)\left(1 - \frac{{w^{min}}^2}{w_i^2}\right)\}$, there exists a feasible assignment of individual i to an intervention bundle in $\mathcal{M}(\kappa)$ such that their FV spending exceeds τ_i.*

Applying Theorem EC.2 can be thought of as a "post-processing" step where investments in the monetary interventions are adjusted downward for each consumer to the extent possible. After applying this adjustment, the number of unique education intervention levels is still M. However, the number of unique monetary intervention levels could be as many as P. Applying this to the FoodAPS dataset using the MLE-exp method with $\eta_v = 5$, $b_i = 0.26$ for all i, and $M = \infty$ results in a 16% reduction in the total variable cost of the program.

The proof of Theorem EC.2 is given below.

Proof of Theorem EC.2. This proof closely follows that of Theorem 2 and Corollary 2. Recall that w_i comes from considering a specific lower bound of the function $h_i(\hat{\alpha})$ for $\hat{\alpha} \geq \hat{\alpha}_i^*$. This lower bound is specifically given by $h_i^*(U) - \frac{l(u_i^*)}{2}(\hat{\alpha} - \hat{\alpha}_i^*(U))^2 \triangleq \underline{h}_i(\hat{\alpha}; U)$ (see the proof of Proposition 4).

By Corollary EC.1, for any U_i' such that $\hat{\alpha}_i^*(U_i') \leq U_i'$, $\hat{\alpha}_i^*(U_i') = \hat{\alpha}_i^*(U)$. Thus we force $U_i \geq \hat{\alpha}_i^*$ so that this property holds. Because the second term of $\underline{h}_i(\hat{\alpha}; U)$, $\frac{l(u_i^*)}{2}(\hat{\alpha} - \alpha_i^*(U))^2$, is invariant to U when $f_{u_i}(u)$ is exponential, $\underline{h}_i(\hat{\alpha}; U_i)$ is identical to $\underline{h}_i(\hat{\alpha}; U)$ on the domain $\hat{\alpha} \geq \alpha_i^*(U_i)$ up to to a vertical shift of $h_i^*(U) - h_i^*(U_i)$.

Our first claim is the following: Let $w_i(U_i)$ be defined analogously to w_i in Theorem 2, so $w_i(U_i) \triangleq \sqrt{\frac{2(h_i^*(U_i)-\tau_i)}{l(u_i^*)}}$. As long as $w_i(U_i) \geq w^{min}$, the construction given in Theorem 2 remains feasible, ensuring that there exists an intervention such that each individual's FV spending exceeds τ_i. This proof of this claim proceeds almost identically to the proof of Theorem 2, with the exception that U_i is substituted for U.

Because $h_i^*(U_i)$ is decreasing as U_i decreases, our goal is to make U_i as small as possible while ensuring that $w_i(U_i) \geq w^{min}$. We proceed by assuming that $w_i(U)$ is already known (since we have the construction from Theorem 2), but $w_i(U_i)$ is not known because $h_i^*(U_i)$ is unknown.

Therefore, we will lower bound $w_i(U_i)$ by lower bounding $h_i^*(U_i)$. A valid lower bound of $h_i^*(U_i)$ is given by $h_i^*(U) - (U - U_i)$. This lower bound is achieved if the entire decrease in individual i's monetary intervention amount—going from $U - \hat{\alpha}_i^*(U)$ to $U_i - \hat{\alpha}_i^*(U)$—translates into a decrease in their FV spending (with OF spending remaining the same). Thus, replacing $h_i^*(U_i)$ with $h_i^*(U) - (U - U_i)$, we get

$$w_i(U_i) \geq \sqrt{\frac{2(h_i^*(U) - (U - U_i) - \tau_i)}{l(u_i^*)}} = \sqrt{\frac{2((h_i^*(U) - \tau_i) - (U - U_i))}{l(u_i^*)}}.$$

Since our goal is to decrease U_i to the extent possible while ensuring that $w_i(U_i) \geq w^{min}$, we wish to solve:

$$\sqrt{\frac{2((h_i^*(U) - \tau_i) - (U - U_i))}{l(u_i^*)}} = w^{min}$$

for $U - U_i$. This results in

$$U - U_i = (h_i^*(U) - \tau_i) - \frac{l(u_i^*) {w^{min}}^2}{2} \tag{EC.16}$$

$$= (h_i^*(U) - \tau_i) - (h_i^*(U) - \tau_i) \frac{{w^{min}}^2}{w_i(U)^2} \tag{EC.17}$$

$$= (h_i^*(U) - \tau_i) \left(1 - \frac{{w^{min}}^2}{w_i(U)^2}\right). \tag{EC.18}$$

The second equality comes from substituting $w_i(U) = \sqrt{\frac{2(h_i^*(U)-\tau_i)}{l(u_i^*)}}$–the definition of $w_i(U)$ given in Theorem 2. □

EC.3.3. Discussion on practicality

While the preceding results provide guidance on the tradeoff between fixed and variable costs, implementing these methods requires intimate knowledge of the individual-level utility functions. This is also true of Theorem 2 in the main text. However, Section 4 showed that by using Theorem 2 to provide guidance on the number of bundles, strong empirical performance could be obtained without knowing the individual utility functions *a priori*.

The same may not be true of the methods proposed in this section. In practice when the true utility functions are unknown, it is more difficult to implement varying-cost bundles relative to equal-cost bundles. As shown in Figure EC.1, reducing someone's total investment amount too much could result in there being no intervention bundle that can achieve FV spending over τ_i. Intuitively, more constrained budgets necessitate the need for more precise, personalized intervention bundles. This is already challenging when total intervention amounts are fixed, and this challenge is exacerbating when intervention amounts are allowed to vary. Therefore, one additional motivation for assuming that everyone receives a total investment amount of U, rather than considering decreasing this amount for certain individuals, is to help improve the robustness of the results.

EC.4. Empirical section details

This section provides more detail on Section 4. Sections EC.4.1 and EC.4.2 provide more details on the data, and in particular how B_i and v_i are computed. Section EC.4.3 provides additional details on the utility function estimate procedures.

EC.4.1. Food expenditures and budget

It is assumed that purchasing decisions during the FoodAPS survey week are representative of each household's average behavior. Each household's average FV and OF spending per grocery shopping trip is taken to be the sum of their FV and OF purchasing, divided by the number of grocery shopping trips, during their survey week. In order to satisfy Assumption 3, we filter the FoodAPS data to SNAP-receiving households that purchased a non-zero amount of both FVs and OF during their survey week. There are 981 households in this population. Furthermore, each household's food budget per grocery trip, B_i, is assumed to be the sum of their grocery expenditures during the survey week divided by the number of grocery shopping trips. Thus, $B_i = h_i + u_i$ by design. Figure EC.4 (left) shows the average FV and OF expenditures per household in the FoodAPS dataset.

EC.4.2. Value of nutrition

Each household's value of nutrition, v_i or v_{i_0}, is estimated based on their response to the binary questions in Table EC.4.

Table EC.4 Survey questions used to construct value of nutrition.

Description
Whether household searched for nutritional information online in the past two months
Whether household participated in a nutrition education event in the past two months
Whether the household uses Nutrition Facts panels always or often
Whether household tried to follow MyPlate OR MyPyramid recommendations
Whether the household has a garden OR receives produce from someone else's garden
Whether the household ever gets food from a farm stand or farmers' market

Each household receives one point per affirmative answer to the questions above. Let the sum of the points for household i be denoted by n_i. A household's value of nutrition is then estimated as

$$\tilde{v}_i = \frac{\min\{\max\{1, n_i\}, 3\} - 1}{2}.$$

Notice that each score is first truncated below by one, and above by 3. The purpose of this truncation is to eliminate tails, since they may not be meaningful. The scores are then normalized to be between zero and one. This provides us with a relative estimate of value of nutrition. Figure EC.4 (right) shows the estimated value of nutrition across the FoodAPS dataset using this method.

While we consider this to be the preferred method for estimating the value of nutrition, there are potential pitfalls in relying solely on survey data. For instance, households might be tempted to exaggerate their nutrition-related behavior because they know they are being surveyed.

Here we discuss possible alternative methods for estimating v_i that do not rely on survey data. First, if the functions $f_{h_i}(h)$ are assumed to lie within a restricted class of functions, it is possible that the chosen functional form allows for the simultaneous identification of v_i and the parameters of $f_{h_i}(h)$. Because v_i) multiplies the function $f_{h_i}(h)$ in the objective function of Problem 1, identification of v_i could only be possible if the assumed functional form of $f_{h_i}(h)$ does not contain an unconstrained scaling parameter. For more flexible functional forms, it is unlikely that this approach would work.

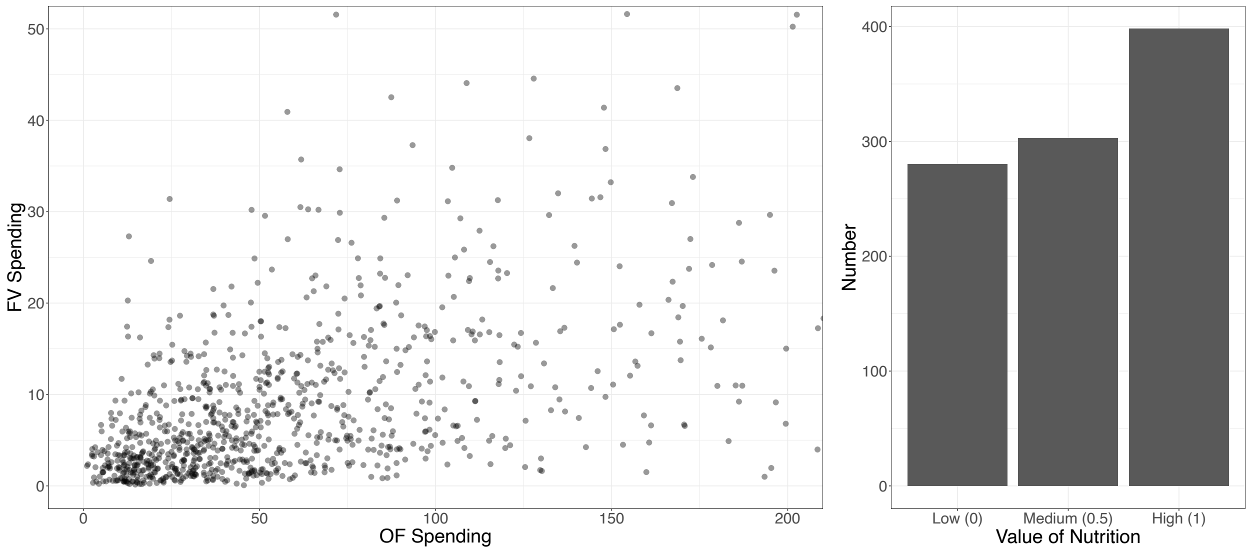

Figure EC.4 **FV and OF purchasing (left), and estimated value of nutrition (right) for the low-income FoodAPS dataset.**

Another potential strategy is the following. Policymakers could start by presuming that all households share a "typical" and homogeneous preference for FVs, resulting in identical functions $f_{h_i}(h)$. With $f_{h_i}(h)$ established, v_i can be deduced from observed grocery shopping habits, and households can then be grouped into interventions based on these inferred parameters. After several months, any shifts in FV and OF expenditures can be monitored and contrasted with anticipated changes. Using this comparison, policymakers can refine their estimates of $f_{h_i}(h)$ and v_i and reassign households to appropriate intervention bundles.

For instance, consider a household with a high value of nutrition but unique aversion for FVs, leading them to buy few FVs. Initially, this household would be perceived as having an average preference for FVs and, to justify their low FV expenditure, a presumed low value for nutrition. However, after a few months, the policymaker might observe minimal changes in their FV spending post-enrollment in a nutrition education program. This lack of change suggests that they already placed a high value on nutrition. The policymaker could then adjust their estimates for $f_{h_i}(h)$ and v_i and reallocate the household to an intervention bundle with a greater allocation to the on monetary intervention.

EC.4.3. Utility function estimation and assignment methods

In this section we provide more details on each method used to generate and estimate the individual-level utility functions.

EC.4.3.1. Sampling and ground truth For a given utility function class parameterized by a vector of length 2 or more, and given individual-level shopping behavior, there are an infinite number of parameters that can explain an individual's observed behavior. This is because each individual's KKT condition is simply given by:

$$(1+v_i)f'_{h_i}(h_i) = f'_{u_i}(u_i) \tag{EC.19}$$

Any set of parameters that satisfies the above equality can rationalize the observed values of h_i and u_i. Let $\mathcal{W}$ be the domain of reasonable parameter values for $f_{h_i}(h)$ and $f_{u_i}(u)$. For example, $\mathcal{W}$ may implicitly impose the constraint that $f_{h_i}(h)$ is more concave than $f_{u_i}(u)$. A feasible set of parameters, $\mathbf{w}_i$, for individual i is one such that EC.19 holds and $\mathbf{w}_i \in \mathcal{W}$.

Suppose that, in aggregate, there are n parameters that parameterize $f_{h_i}(h)$ and $f_{u_i}(u)$. Thus, $\mathbf{w}_i \in \mathcal{W} \subset \mathbb{R}^n$. The Sampling procedure proceeds along the following steps.

1. Randomly generate values of $\mathbf{w}_{1,\dots,n-1}$ by sampling from the first $n-1$ dimensions of $\mathcal{W}$.
2. Next, solve Equation EC.19 in order to obtain w_n.
3. If $\mathbf{w} \in \mathcal{W}$, stop. Otherwise, repeat steps 1-2 until $\mathbf{w} \in \mathcal{W}$.

This method can be used to either generate plausible ground truth parameters, or to sample feasible individual utility functions from the specified domain. Thus, the same process is used in for both the Sampling method used in the main text, and to generate ground truth functions that are used to evaluate the proposed assignment methods.

For exponential utility functions, we define $\mathcal{W} = \{\eta_{u_i},\ \eta_{u_i},\ c_{u_i},\ c_{h_i} : \ \eta_{u_i} \in [0.005, 0.06],\ \eta_{h_i} \in [0.005, 0.06],\ c_{u_i} \in [0.1, 40],\ c_{h_i} = 1,\ \eta_{h_i} \geq \eta_{u_i}\}$ in the numerical experiments. For isoelastic utility functions, we define $\mathcal{W} = \{\eta_{u_i},\ \eta_{u_i},\ c_{u_i},\ c_{h_i} : \eta_{u_i} \in [0.1, 0.5],\ \eta_{h_i} \in [0.1, 0.95],\ c_{u_i} \in [0.1, 5],\ c_{h_i} = 1,\ \eta_{h_i} \geq \eta_{u_i}\}$ in the numerical experiments. We note that setting $c_{h_i} = 1$ is without loss of generality.

Figure EC.4.3.1 shows an example of twenty plausible "ground truth" utility functions for a households—ten assuming an isoelastic utility function class and ten assuming an exponential utility function class. All pairs of utility functions rationalize the observed data point for this household.

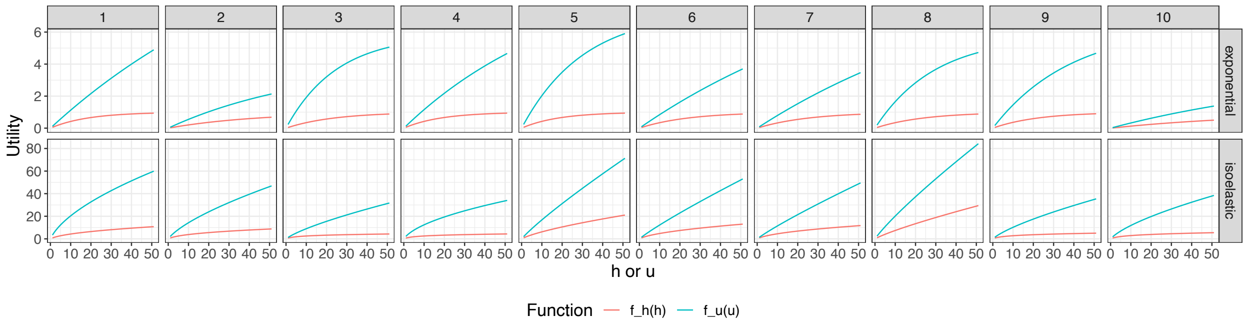

Figure EC.5 **Ten instances of ground truth utility functions for one individual and for both the exponential (top)and isoelastic (bottom) class of utility functions.**

EC.4.3.2. Mixed-effects MLE estimation In the approach above, each individual's utility functions were independent of one another. Under this independence assumption, for each individual we only had one data point to use for estimation and thus we could not fully identify the individual parameters. However, if we assume that the individual utility functions are related, we can exploit the entire population's data to estimate each individual's parameters, allowing for full identification. This does *not* guarantee that we can recover the ground truth, rather, it is another approach that can be used to estimate the individual utility functions.

In this approach, we employ a mixed-effects model. This assumes that there is a population-level mean for each parameter and that the individuals in the population do not vary substantially from this mean. For example, if the functions $f_{u_i}(\cdot)$ and $f_{h_i}(\cdot)$ are within the class of exponential utility functions, there is an assumed population-level mean for c_{u_i}, c_{h_i}, η_{u_i}, and η_{h_i}. The estimation procedure thus penalizes the variance of c_{u_i}, c_{h_i}, η_{u_i}, and η_{h_i} with respect to this mean.

To identify the parameters of $f_{h_i}(h)$ and $f_{u_i}(u)$, we follow a two-step estimation procedure. This is necessary because of the potential confounding between v_{i_0} and the parameters of $f_{h_i}(h)$. In other words, we need to differentiate between households' utility for FV and their value of nutrition.

Our general procedure is as follows. We will assume that the functions $f_{h_i}(\cdot)$ and $f_{u_i}(\cdot)$ are within a finitely-parameterized function classes with parameter vector $\boldsymbol{\gamma_i}$ for individual i. Furthermore, we employ a mixed-effects model in which we assume that $\boldsymbol{\gamma_i}$ is drawn from a population-level distribution that is finitely-parameterized by $\boldsymbol{\theta}$. Therefore, $\boldsymbol{\gamma}_i \sim f(\boldsymbol{\gamma};\boldsymbol{\theta})$ where $f(\boldsymbol{\gamma};\boldsymbol{\theta})$ is the density function of $\boldsymbol{\gamma_i}$ with parameter vector $\boldsymbol{\theta}$.

The two steps are:

1. Use the population of households for which $v_{i_0} = 0$ to estimate the population-level parameter vector $\boldsymbol{\theta}$.
2. With $\boldsymbol{\theta}$ fixed, estimate the individual-level parameters $\boldsymbol{\gamma_i}$ for each household.

Each stage above follows a maximum-likelihood procedure. However, because h_{i_0} and u_{i_0} are implicitly defined by the consumer's nonlinear KKT conditions, and due to the inclusion of both individual- and population-level parameters, the maximum likelihood problem is not straightforward. In general, with outcome variables $\mathbf{y}$, covariates $\mathbf{x}$, population-level parameters $\boldsymbol{\theta}$ and individual-level (random or latent) effects $\boldsymbol{\gamma}$, maximum likelihood estimation seeks to maximize the marginal data likelihood, namely $f(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$, with respect to $\boldsymbol{\theta}$. This is a natural quantity to wish to maximize since we know the ground truth of the outcome variables. In non-linear mixed effects (or latent variable) models, maximizing the marginal data likelihood if often done using stochastic expectation-maximization (EM) algorithms, which include an "S-step" to approximate taking an expectation over $f(\boldsymbol{\gamma}|\boldsymbol{\theta}, \mathbf{y}, \mathbf{x})$. Instead of taking this expectation directly, the S-step samples from the distribution $f(\boldsymbol{\gamma}|\boldsymbol{\theta}, \mathbf{y}, \mathbf{x})$, which is often accomplished using Markov chain Monte Carlo methods. For more information on stochastic EM, see Nielsen et al. (2000) and Delyon et al. (1999).

For the specific problem presented in this paper, each iteration of stochastic EM requires solving a system of nonlinear equations many times, and is not guaranteed to converge. Therefore, we choose an alternative approach that works well in practice. Instead of seeking to maximize the marginal data likelihood, we maximize the complete likelihood, $f(\mathbf{y}, \boldsymbol{\gamma}|\mathbf{x}, \boldsymbol{\theta})$ with respect to both $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$. The advantage of this approach is that the complete likelihood can be maximized by solving one optimization problem that produces estimates of both $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ simultaneously. This method is known to produce similar estimates to those produced by maximizing the marginal data likelihood, and is much simpler and faster in this context.

In what follows, we describe the estimation procedure for the two stages outlined above in greater detail. The detailed procedure for Stage 1, which estimates the population-level parameters of $f_{u_i}(u)$ and $f_{h_i}(h)$, is shown in Algorithm 1. The function Optimize$(\cdot)$ is used in each stage, and is described below.

$$
\begin{aligned}
\text{Optimize(pTBD, pFixed, obj\_fun, } \mathcal{P}) = \min_{\mathbf{h}, \mathbf{u}, \text{ pTBD}} \; & \text{obj\_fun}(\mathbf{h}, \mathbf{u}, \text{pTBD, pFixed}) \\
s.t. \; & (h_i, u_i) \text{ solves KKT}_i \;\; \forall\, i \in \mathcal{P} \\
& h_i \geq 0, \; u_i \geq 0 \;\; \forall i \in \mathcal{P}
\end{aligned}
$$

where KKT_i are the KKT conditions of consumer i's problem (Equation EC.1 for consumer i). The variable pTBD contains the parameters to be optimized, and pFixed contains the parameters that have already been determined.

Algorithm 1 first subsets the population to those with zero value of nutrition to determine $\bar{c_h}$, $\bar{\eta}_h$, $\bar{c_u}$, and $\bar{\eta}_u$. In Algorithm 1, an asterisk denotes a quantity *after* optimization/maximum likelihood estimation. For example, pTBD* contains the optimized values of these parameters.

Algorithm 1 Stage 1

Input: v_i, h_i, and u_i for $i \in \mathcal{P}$; weight γ

1: $\mathcal{P}' \leftarrow \{i \in \mathcal{P} : v_i = 0\}$ ▷ Subset population

2: pTBD $\leftarrow \{\eta_{h_i}, \eta_{u_i}, c_{h_i}, c_{u_i}\}_{i\in\mathcal{P}'} \cup \{\bar{\eta}_h, \bar{\eta}_u, \bar{c}_h, \bar{c}_u\}$

3: pFixed $\leftarrow \emptyset$

4: obj_fun $\leftarrow \sum_{i\in\mathcal{P}'} \left(\frac{(\tilde{h}_i - h_i)^2}{\bar{h}^2} + \frac{(\tilde{u}_i - u_i)^2}{\bar{u}^2} + \gamma\left((\eta_{h_i} - \bar{\eta}_h)^2 + (\eta_{u_i} - \bar{\eta}_u)^2 + (c_{h_i} - \bar{c}_h)^2 + (c_{u_i} - \bar{c}_u)^2\right)\right)$

5: pTBD* $\leftarrow$ Optimize(pTBD, pFixed, obj_fun, $\mathcal{P}'$)

6: parmsFixed $\leftarrow \bar{\eta}_h^*, \bar{\eta}_u^*, \bar{c}_h^*, \bar{c}_u^*$ ▷ Extract population means from pTBD* and store them in pFixed

Return: pFixed

After Stage 1, $\bar{c_h}$, $\bar{\eta}_h$, $\bar{c_u}$, and $\bar{\eta}_u$ are fixed. Stage 2 then uses the entire population to determine the individual parameters c_{h_i}, η_{h_i}, c_{u_i}, and η_{u_i} for all $i \in \mathcal{P}$. Stage 2 is described in Algorithm 2. Table EC.5 summarizes the distributions of c_{h_i}, η_{h_i}, c_{u_i}, and η_{u_i} across the FoodAPS population.

Algorithm 2 Stage 2

Input: v_i, h_i, and u_i for $i \in \mathcal{P}$; pFixed; weight γ

pTBD $\leftarrow \{\eta_{h_i}, \eta_{u_i}, c_{h_i}, c_{u_i}\}_{i\in\mathcal{P}}$

obj_fun $\leftarrow \sum_{i\in\mathcal{P}} \left(\frac{(\tilde{h}_i - h_i)^2}{\bar{h}^2} + \frac{(\tilde{u}_i - u_i)^2}{\bar{u}^2} + (\eta_{h_i} - \bar{\eta}_h)^2 + \gamma\left((\eta_{u_i} - \bar{\eta}_u)^2 + (c_{h_i} - \bar{c}_h)^2 + (c_{u_i} - \bar{c}_u)^2\right)\right)$

pTBD* $\leftarrow$ Optimize(pTBD, pFixed, obj_fun, $\mathcal{P}$)

pFixed $\leftarrow$ pFixed $\cup \{\eta_{h_i}^*, \eta_{u_i}^*, c_{h_i}^*, c_{u_i}^*\}_{i\in\mathcal{P}}$ ▷ Extract individual-level parameters from pTBD*, save them in pFixed

Return: pFixed

EC.4.4. Determining the number of bundles to consider

We use the Samp-iso, Samp-exp, MLE-iso, and MLE-exp methods to determine a reasonable number of intervention bundles in Section 4.4. From Theorem 2, we must choose κ. We

Table EC.5 Distribution of individual parameters estimated using mixed-effects model following Algorithms 1 and 2 with $\gamma = .01$.

	Min	Mean	Max	StdDev
	Isoelastic functional form			
c_{u_i}	1.35	1.64	2.04	0.0537
η_{u_i}	0.100	0.206	0.554	0.0731
η_{h_i}	0.100	0.358	0.732	0.0663
	Exponential functional form			
c_{u_i}	2.05	3.21	5.00	0.538
η_{u_i}	0.0050	0.00662	0.0117	0.00118
η_{h_i}	0.0050	0.0137	0.0500	0.0106

use $\kappa = w^{min} \triangleq \min_{i \in P} w_i$, where $w_i = \sqrt{\frac{2(h_i^* - \tau_i)}{l(u_i^*)}}$ and $l(u^*) = b\frac{A_{f_{u_i}}(u_i^*)A_v^+}{A_{f_{h_i}}^-} + b^2 A_{f'_{u_i}}^+$. Here we are using the shorthand $A_f(y) \triangleq \frac{-f''(y)}{f'(y)}$, $A_f^+ \triangleq \max_y \frac{-f''(y)}{f'(y)}$, and $A_f^- \triangleq \min_y \frac{-f''(y)}{f'(y)}$.

For the function $v_i(\alpha)$ given in Equation 6, $A_v^+ = \eta_v / U$ for all i. When utility functions are exponential, $A_{f_{u_i}}(u_i^*) = A_{f'_{u_i}}^+ = \eta_{u_i}$ and $A_{f_{h_i}}^- = \eta_{h_i}$. Therefore, in this case w_i reduces to

$$\sqrt{\frac{2(h_i^* - \tau_i)}{\eta_{u_i} b_i \frac{\eta_v}{U} \frac{1}{\eta_{h_i}} + b_i^2 \eta_{u_i}}}.$$

When utility functions are isoelastic, $A_{f_{u_i}}(u_i^*) = \frac{\eta_{u_i}}{1+u_i^*}$, $A_{f'_{u_i}}^+ = \frac{1+\eta_{u_i}}{1+\min_{\hat{\alpha}\in[0,U]} u_i(\hat{\alpha})}$ and $A_{f_{h_i}}^- = \frac{\eta_{h_i}}{1+h_i^*}$. Recall that $u_i(\hat{\alpha})$ is a decreasing function, and is thus minimized when $\hat{\alpha} = U$. Therefore, a lower bound on w_i is given by

$$\sqrt{\frac{2(h_i^* - \tau_i)}{\frac{b_i \eta_v}{U} \frac{\eta_{u_i}}{\eta_{u_i}} \frac{1+h_i^*}{1+u_i(U)} + b^2 \frac{1+\eta_{u_i}}{1+u_i(U)}}}.$$

To determine w^{min}, we take the empirical minimum of these quantities over the FoodAPS population. Figure 3 shows the resulting number of bundles required after taking the empirical minimum of these quantities as τ_i varies.

EC.5. Additional empirical results

Figures EC.6 - EC.10 and Table EC.5 show additional computation results under varying pairs of linking functions. We show results for $b \in \{0.26, 0.36\}$ and $\eta_v \in \{1, 2, 5\}$. When $\eta_v = 1$, $v(\alpha)$ is near-linear. When $\eta_v = 5$, $v(\alpha)$ is more concave. Table EC.5 shows the results of the one-size-fits-all methods considered. While Best OSFA does not vary as b_i increase for a fixed value of η_v, it does vary as η_v changes. When $\eta_v = 1$, Best OSFA consists of a 50% investment in education, decreasing to 40% for $\eta_v = 2$ and 30% for $\eta_v = 5$. Intuitively, the reason for this is because when η_v increases, the effect of the education intervention is steep at first but has very diminishing returns. Therefore, investing a smaller portion of the budget into education is better. Edu only continues to perform poorly across all linking functions considered. The performance of Price only decreases as η_v increases, and improves as b_i increases. In contrast, the performance of Best OSFA improves as η_v increases.

Figures EC.6 - EC.10 show the performance of the personalizable methods across the linking functions considered as M increases. Overall, Samp-iso and MLE-iso exhibit robust performance across all linking functions considered. They consistently perform well, and generally improve as M increases. In contrast, Samp-exp appears to "over-fit," with performance often declining as M increases. Additionally, for the linking functions considered in this section as opposed to the main text, the performance of Samp-exp is worse relative to the other methods. The performance of MLE-exp lies in the middle—while it does not have the best performance in general, its performance is not as poor as MLE-exp.

Table EC.6 Quantiles of the optimality ratio for one-size-fits-all methods.

	$\eta_v = 1,\ b = 0.26$						$\eta_v = 1,\ b = 0.36$				
Method	Min	1%	5%	10%	25%		Min	1%	5%	10%	25%
Edu only	0	0	0	0	0		0	0	0	0	0
Price only	0	0.02	0.05	0.09	0.24		0	0.03	0.07	0.13	0.34
Best OSFA (50)	0.47	0.5	0.5	0.5	0.5	(50)	0.48	0.5	0.5	0.5	0.5
	$\eta_v = 2,\ b = 0.26$						$\eta_v = 2,\ b = 0.36$				
Method	Min	1%	5%	10%	25%		Min	1%	5%	10%	25%
Edu only	0	0	0	0	0		0	0	0	0	0
Price only	0	0.01	0.04	0.07	0.18		0	0.02	0.05	0.1	0.25
Best OSFA (40)	0.49	0.6	0.6	0.6	0.6	(40)	0.5	0.6	0.6	0.6	0.6
	$\eta_v = 5,\ b = 0.26$						$\eta_v = 5,\ b = 0.36$				
Method	Min	1%	5%	10%	25%		Min	1%	5%	10%	25%
Edu only	0	0	0	0	0		0	0	0	0	0
Price only	0	0.01	0.03	0.06	0.15		0	0.02	0.04	0.08	0.21
Best OSFA (30)	0.67	0.7	0.7	0.7	0.7	(30)	0.69	0.7	0.7	0.7	0.7

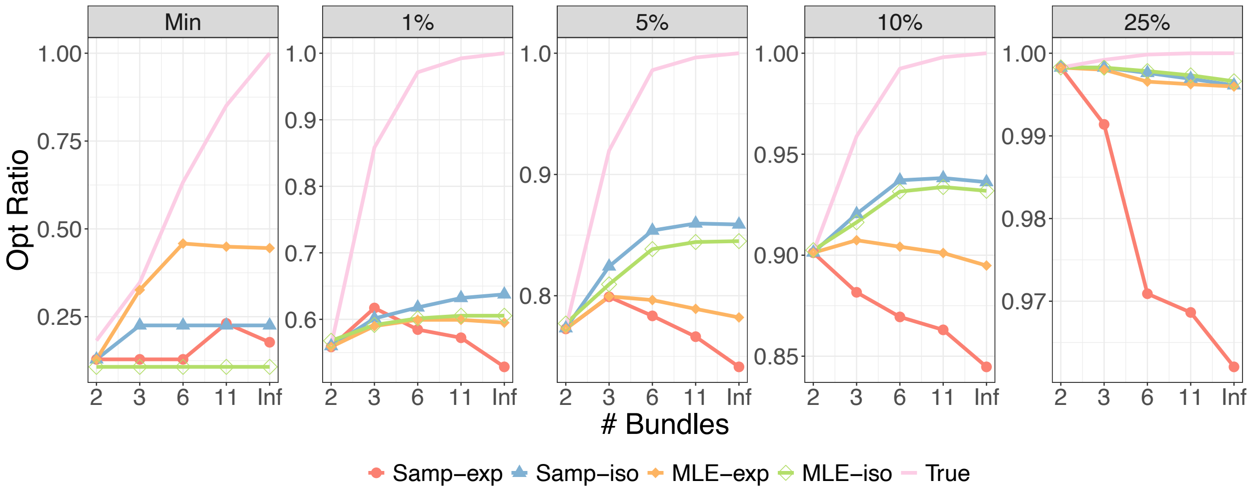

Figure EC.6 Performance of personalizable methods and True benchmark when $\eta_v = 2$ and $b = 0.26$.

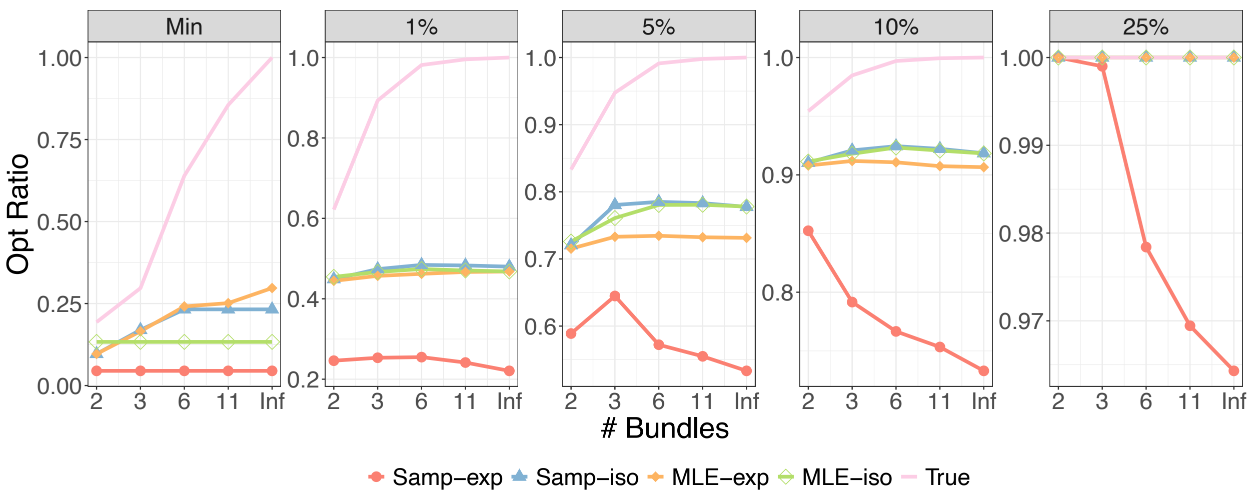

Figure EC.7 Performance of personalizable methods and True benchmark when $\eta_v = 1$ and $b = 0.26$.

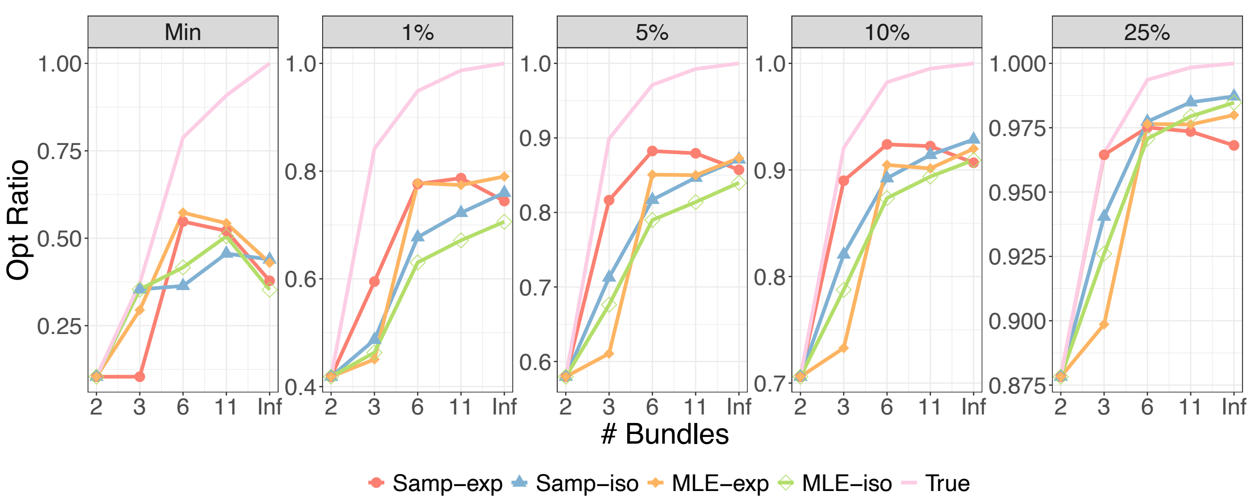

Figure EC.8 Performance of personalizable methods and True benchmark when $\eta_v = 5$ and $b = 0.36$.

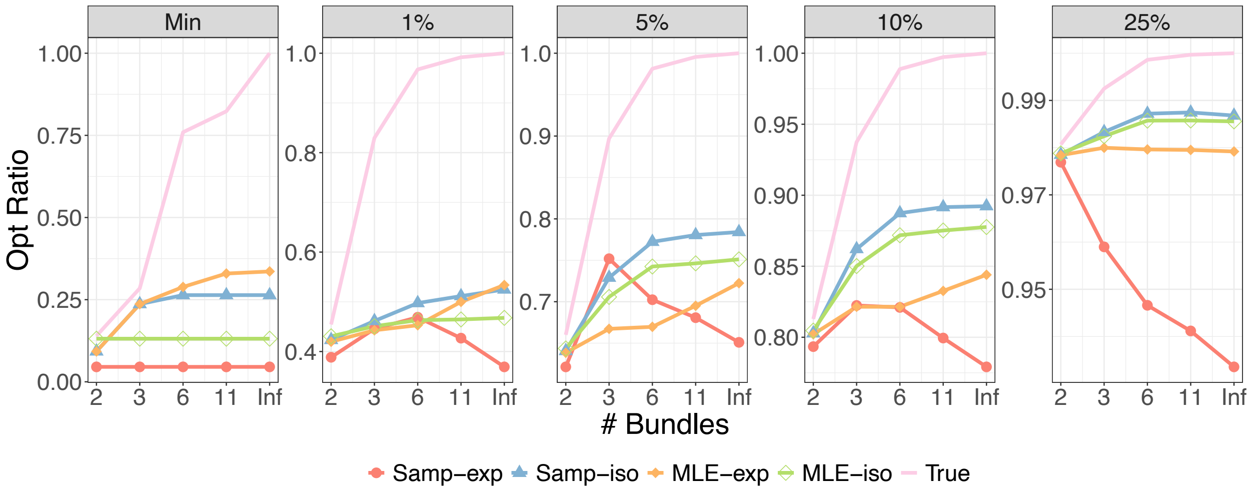

Figure EC.9 Performance of personalizable methods and True benchmark when $\eta_v = 2$ and $b = 0.36$.

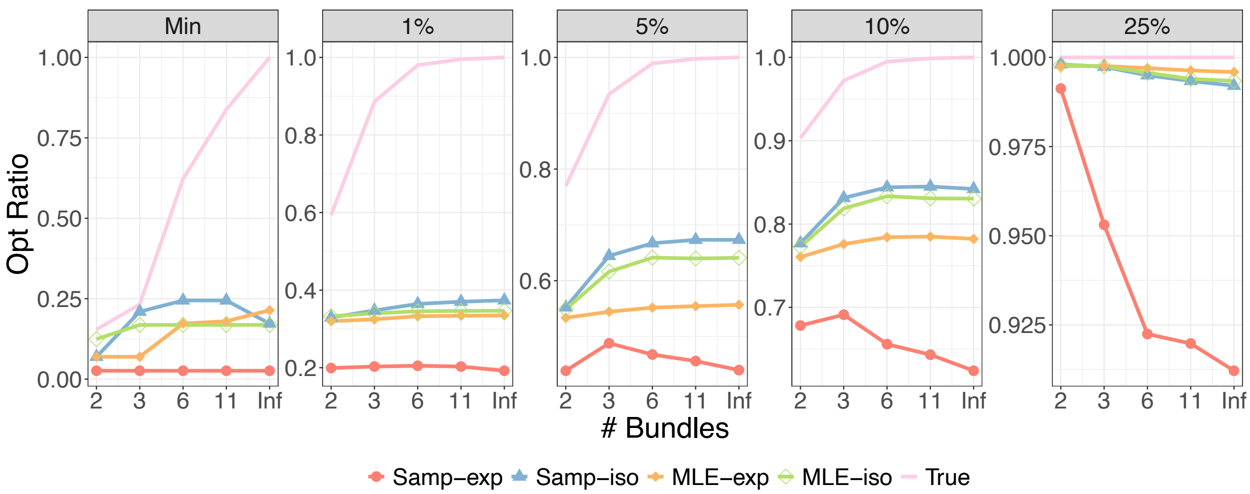

Figure EC.10 Performance of personalizable methods and True benchmark when $\eta_v = 1$ and $b = 0.36$.

EC.6. Alternative assignment methods

This section considers four additional assignment methods. The first two compute the maximum (and minimum) investment in education that could be optimal for each household. The last two assignment methods only differentiate between households based on their value of nutrition. These methods may be desirable for their simplicity and/or fairness properties.

EC.6.1. Maximal assignments

This section considers two additional assignment methods, described below:

1. Price-max. For each individual, this method finds the set of parameters across the exponential and isoelastic function classes that maximizes the investment in the mon-

etary intervention. Additional functional classes could be readily included as well. In other words, we ask the question: *what is the maximum investment in the monetary intervention that could be optimal for individual i, given their observed behavior?* The motivation for this method stems from the reasonable assumption that all individuals would prefer the monetary intervention to the educational intervention. Therefore, a fair approach to administering interventions might be to give each individual the maximum possible monetary intervention that is reasonable given their observed shopping behavior.

2. Edu-max. This approach is analogous to Price-max, but maximizes the investment in education.

To compute the Price-max assignment, we follow Algorithm 3, described below, for each individual in the population. Algorithm 3 loops through the utility function classes considered and finds the set of parameters within a specified domain that explain the individual's observed behavior *and* maximize the investment in the monetary intervention. An analogous algorithm is used to compute Edu-max.

Algorithm 3 Price-Max Estimation

Input: $B_i, v_i, h_{0_i}, u_{0_i}$ ▷ Data for individual i

Input: $\mathcal{C}$ ▷ Utility function families considered

Input: tol ▷ Tolerance for estimation error

Input: $\mathbf{W}_c$ for $c \in \mathcal{C}$ ▷ Parameter domain for utility function class c

for $c \in \mathcal{C}$ **do** ▷ Iterate through utility function classes

$$\beta_c = \arg \max_{\substack{\beta\in[0,U],\\ \mathbf{w}\in\mathbf{W}_c}} \beta$$

s.t. $|h_i^*(0,0;\mathbf{w}) - h_{0_i}| \leq tol$ ▷ Ensure parameters explain observed shopping behavior

$|u_i^*(0,0;\mathbf{w}) - u_{0_i}| \leq tol$

$\alpha = U - \beta$

$(\beta,\, \alpha,\, h_i^*(\beta,\, \alpha;\mathbf{w}),\, u_i^*(\beta,\, \alpha;\mathbf{w}))$ satisfies Proposition 3 ▷ Ensure optimality

end for

$\beta \leftarrow \max_{c\in\mathcal{C}}\{\beta_c\}$

Return: β ▷ Output the robust estimate of β

To build intuition for these methods, Figure EC.11 reproduces Figure 4 with these two methods included. As expected, Price-max always results in the lowest value of $\hat{\alpha}_i^*$ (i.e.,

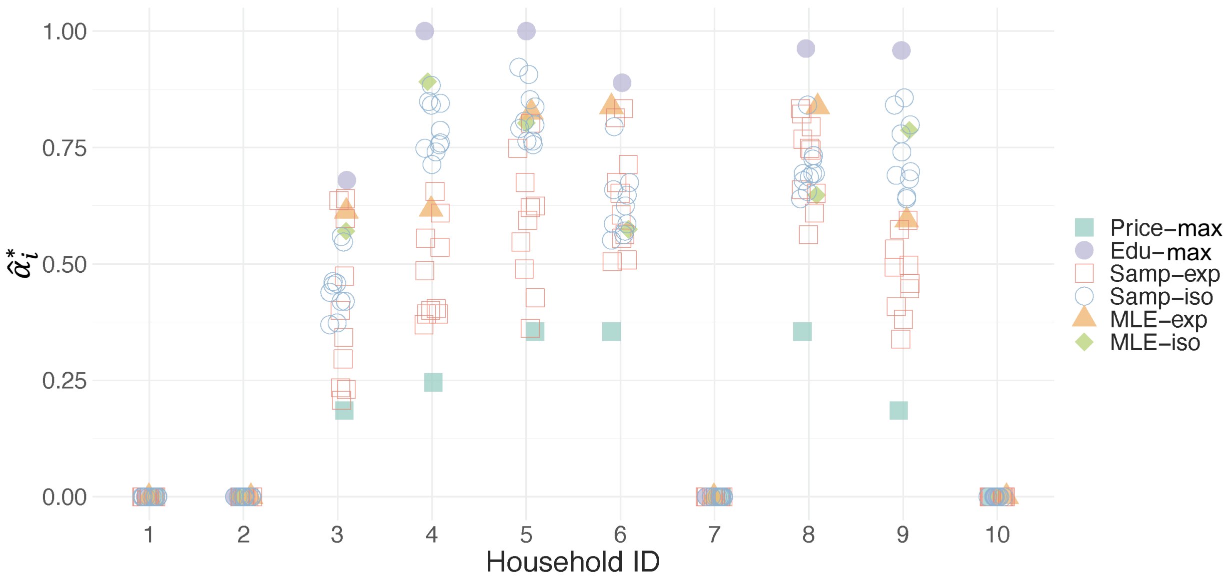

Figure EC.11 **Estimated value of $\hat{\alpha}_i^*$ for 10 randomly selected households.**

the highest investment in the monetary intervention), and Edu-max always results in the highest value of $\hat{\alpha}_i^*$.

Table EC.7 shows the performance of Price-max and Edu-max compared to the methods considered in the main text. Although these methods may be desirable for certain reasons (e.g., fairness concerns), based purely on performance they appear to be inferior to the other methods considered.

	3 Bundles					*6 Bundles*				
Method	Min	1%	5%	10%	25%	Min	1%	5%	10%	25%
Price-max	0.02	0.06	0.14	0.22	0.68	0.09	0.63	0.73	0.78	0.9
Edu-max	0.17	0.56	0.77	0.86	0.95	0.17	0.64	0.80	0.88	0.96
Samp-exp	0.14	0.63	0.83	0.9	0.97	0.45	0.81	0.91	0.94	0.98
Samp-iso	0.44	0.58	0.77	0.86	0.96	0.47	0.73	0.85	0.92	0.98
MLE-exp	0.46	0.56	0.71	0.82	0.94	0.63	0.74	0.83	0.89	0.97
MLE-iso	0.46	0.57	0.76	0.85	0.95	0.48	0.68	0.83	0.9	0.98

Table EC.7 **Performance of the robust assignment procedures.**

EC.6.2. Value of nutrition only

In this section we consider two methods that only differentiate between households based on v_{i_0}. To motivate this approach, Figure EC.6.2 shows the importance of each consumer-level variable in terms of determining $\hat{\alpha}_i^*$ for that method. Figure EC.6.2 (left) shows the variable importance derived from a permutation test. In this test, each variable is randomly

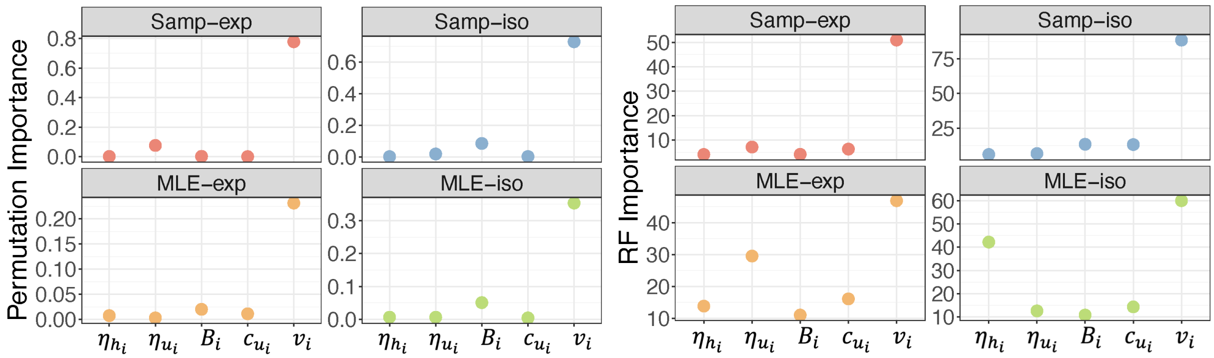

Figure EC.12 Importance of variables in terms of determining $\hat{\alpha}_i^*$, using a permutation test (left) and the Random Forest variable importance measure (right). The permutation importance test uses linear regression with all two-way interactions.

permuted, and the accuracy resulting from a model trained with that variable permuted is compared to the accuracy of the model with all (unpermuted) variables. A larger decrease in accuracy indicates a more important variable. For this test, we use a linear regression with all two-way interactions as the underlying model, and use R^2 as the accuracy metric. Figure EC.6.2 (right) shows the variable importance derived from a random forest model using the increase in node purity as the importance metric. Based on both importance metrics, we see that value of nutrition is by far the most important variable in determining $\hat{\alpha}_i^*$ across all four methods. (We note that η_{h_i} and η_{u_i} always refer to the shape parameter of either the exponential or isoelastic utility functions, and c_{u_i} always refers to the scale parameter of f_{u_i}.) This begs the question as to whether an intervention assignment based *almost entirely* on value of nutrition could perform well.

We therefore consider two variations of assignment methods that differentiate between individuals based only on v_{i_0}. In the first, we assume isoelastic utility functions and that everyone has the median values of the parameters estimated by MLE-iso, as well as the median value of B_{i_0}, so that the only variation at the individual-level is v_{i_0}. This ensures that the estimation of $\hat{\alpha}_i^*$, and the assignment of individuals to bundles, depends only on v_{i_0}. This results in the following policy: assign everyone with a low value of nutrition to the bundle $(0.88U, 0.12U)$, everyone with a medium value of nutrition to $(0.7U, 0.3U)$, and everyone with a high value of nutrition to $(0, U)$. We will refer to this method as MLE-iso-v-only.

The next method is analogous but uses MLE-exp instead (and assumes exponential utility functions). This method results in the following policy: assign those with low value of

	3 Bundles					*6 Bundles*				
Method	Min	1%	5%	10%	25%	Min	1%	5%	10%	25%
MLE-exp-v-only	0.00	0.02	0.11	0.46	0.97	0.00	0.02	0.11	0.46	0.97
MLE-iso-v-only	0.28	0.75	0.85	0.90	0.98	0.28	0.75	0.85	0.90	0.98
Samp-exp	0.14	0.63	0.83	0.9	0.97	0.45	0.81	0.91	0.94	0.98
Samp-iso	0.44	0.58	0.77	0.86	0.96	0.47	0.73	0.85	0.92	0.98
MLE-exp	0.46	0.56	0.71	0.82	0.94	0.63	0.74	0.83	0.89	0.97
MLE-iso	0.46	0.57	0.76	0.85	0.95	0.48	0.68	0.83	0.9	0.98

Table EC.8 Performance of "v-only" approaches.

nutrition to $(0.78U, 0.28U)$, those with a medium value of nutrition to $(0.61U, 0.39U)$, and those with a high value of nutrition to $(0, U)$. We will refer to this method as MLE-exp-v-only.

Both MLE-iso-v-only and MLE-exp-v-only always use three bundles, since there are three estimated levels for value of nutrition. Table EC.8 compares the performance of these methods to the methods suggested in the main text. We compare the performance when $M = 3$ for all methods, and when $M = 6$ (for the methods suggested in the main text and the True benchmark). With the exception of the minimum, MLE-iso-v-only performs quite well at all percentiles. When all methods use three bundles, MLE-iso-v-only is the best method for 99% of the population. When the original methods use six bundles, the performance of MLE-iso-v-only remains comparable to the other methods. By contrast, MLE-exp-v-only struggles to achieve good performance for low percentiles.

This gap in performance between MLE-exp-v-only and MLE-iso-v-only is interesting given that the two methods are similar at first glance. The estimate of v_{i_0} is the same across all methods, and thus the methods only differ in their chosen bundles. MLE-iso-v-only provides both medium and low-value of nutrition households with bundles that have a higher investment in education than MLE-exp-v-only. This seemingly small difference creates a huge impact on downstream FV spending. This suggests that it is not obvious how to choose population-level intervention bundles even in a simple setting where assignments only depend on v_{i_0}.

EC.7. Robustness to misestimation and misspecification

In this section we perform two robustness checks. First, we explore how much the performance of the proposed methods degrades when value of nutrition is misestimated. Second, we consider a situation in which the linking functions are misspecified. In both cases, we find that the proposed methods are robust at the 5th and higher percentiles, though there is degradation at the minimum and first percentiles. We also find that MLE-iso-v-only is a relatively robust method.

EC.7.1. Impact of misestimated value of nutrition

In this section we consider how the performance of our proposed methods changes if we assume that v_{i_0} contains estimation error. In the main text, we assumed that v_{i_0} is estimated precisely.

Let v_{i_0} now denote someone's *true* value of nutrition, and let $\hat{v}_{i_0}$ denote their estimated value of nutrition. We will assume that $\hat{v}_{i_0} = 1$ implies $v_{i_0} = 1$ and $\hat{v}_{i_0} = 0$ implies $v_{i_0} = 0$. Thus, people with high and low value of nutrition, respectively, can be detected. However, for those with a medium value of nutrition (i.e., $\hat{v}_{i_0} = 0.5$), we may have estimation error and $v_{i_0} \neq \hat{v_{i_0}}$. For our assessment, we will generate "true" value of nutrition by assuming that for those with $\hat{v}_{i_0} = 0.5$, $v_{i_0} \sim \text{Normal}(0.5, 0.2)$. The planner only has access to $\hat{v}_{i_0}$.

Table EC.7.1 shows the results of the methods from the main text, in addition to the v-only methods. At the minimum and first percentiles, the performance of Samp-exp, Samp-iso, MLE-exp , and MLE-iso in Table EC.7.1 is significantly worse than in Table 2. However, at the 5th, 10th, and 25th percentiles their performance is comparable to the setting without estimation error. Therefore, estimation error can harm performance for certain individuals, but for most personalization it still superior to the best one-size-fits-all strategies.

Comparing these results to Table EC.8, we see that the performance of MLE-exp-v-only remains consistent, though its performance was not strong to begin with. The performance of MLE-iso-v-only also remains consistent with the exception of the first percentile in which we observe a decrease from 0.75 to 0.57. Therefore, MLE-iso-v-only is relatively robust to misestimation in v_{i_0}.

EC.7.2. Impact of imprecise linking functions

In addition to understanding the consumer-level characteristics, the performance of the proposed methods also relies on knowledge about the linking functions—the effect of the monetary and education intervention on a consumer's food budget and value of nutrition, respectively. While accurate information on intervention effectiveness would be useful even for a one-size-fits-all program, it is particularly important when seeking to personalize interventions.

	3 Bundles						*6 Bundles*				
Method	Min	1%	5%	10%	25%		Min	1%	5%	10%	25%
Edu only	0	0	0	0	0		0	0	0	0	0
Price only	0	0.01	0.03	0.06	0.15		0	0.01	0.03	0.06	0.15
Best OSFA (30)	0.67	0.7	0.7	0.7	0.7	(30)	0.67	0.7	0.7	0.7	0.7
MLE-exp-v-only	0	0.02	0.11	0.41	0.97		0	0.02	0.11	0.41	0.97
MLE-iso-v-only	0.29	0.57	0.82	0.89	0.97		0.29	0.57	0.82	0.89	0.97
Samp-exp	0	0.51	0.79	0.89	0.96		0.2	0.7	0.89	0.93	0.98
Samp-iso	0	0.42	0.74	0.84	0.95		0	0.58	0.83	0.9	0.98
MLE-exp	0	0.33	0.67	0.8	0.94		0.2	0.57	0.8	0.87	0.96
MLE-iso	0	0.37	0.71	0.83	0.95		0	0.51	0.8	0.89	0.98

Table EC.9 Performance when v_{i_0} is misestimated for those with a medium estimated value of nutrition. Results shown for $\eta_v = 5$ and $b = 0.26$.

For many programs that are widely adopted, program effectiveness is measured via small-scale pilots, with randomized control trials being the gold standard. However, such impact evaluations often focus on the average treatment effect (and, even if heterogeneous treatment effects are estimated, they may be inaccurate or imprecise).

For example, if intervention bundles are assigned based on the an average treatment effect, it is possible that individuals will be incorrectly assigned to an intervention bundle relative to their assignment if their true heterogeneous treatment effect were known.

In the following experiment, we generate heterogeneous ground truth effects, given by $\eta_{v_i} \sim \text{Normal}(5, 1.66^2)$ and $b_i \sim \text{Normal}(0.26, .052^2)$ (if negative values arise, we take the maximum with 0.01 for η_v and with zero for b). This results in a significant level of variance. Under the specified distribution, η_{v_i} ranges from $[0.01, 10.4]$ and $b_i \in [0.098, 0.43]$.

To assess the impact of misspecified linking functions, we will look at the performance of the methods from the main text (i.e., where $\hat{\alpha}_i^*$ is determined based on a population-level, average treatment effect with $b_i = 0.26$ for all i, and $\eta_{v_i} = 5$ for all i), and assess the populations' FV spending under their assigned intervention bundle when their ground truth linking functions are heterogeneous, specified by the above distributions.

Table EC.7.2 shows the results for $M = 3$ and $M = 6$. In both cases, the minimum values in Table EC.7.2 are much smaller than that of Table 2. Therefore, there are some individuals for which knowledge of their exact heterogeneous effect is crucial, and assuming an average effect for them does not work well. However, at all percentiles except the minimum, the results of EC.7.2 are very similar to that of Table 2.

When $M = 3$, all personalized methods perform better than Best OSFA for at least 95% of the population, and for 99% of the population when $M = 6$. Therefore, even when the linking functions are misspecified, there is value in personalization. Interestingly, the minimum optimality ratio obtained by MLE-iso-v-only only decreases by 0.04 in this setting, compared to Table EC.8. Because this approach personalizes only on the basis of v_{i_0}, it appears to be robust to misspecifications.

	3 Bundles						*6 Bundles*				
Method	Min	1%	5%	10%	25%		Min	1%	5%	10%	25%
Edu only	0	0	0	0	0		0	0	0	0	0
Price only	0	0.01	0.03	0.06	0.15		0	0.01	0.03	0.06	0.15
Best OSFA (20)	0.48	0.57	0.64	0.67	0.72	(20)	0.48	0.57	0.64	0.67	0.72
MLE-exp-v-only	0	0.02	0.11	0.5	0.97		0	0.02	0.11	0.5	0.97
MLE-iso-v-only	0.24	0.7	0.84	0.9	0.98		0.24	0.7	0.84	0.9	0.98
Samp-exp	0.01	0.62	0.82	0.90	0.97		0.04	0.77	0.89	0.93	0.98
Samp-iso	0.01	0.55	0.77	0.86	0.96		0.02	0.7	0.85	0.92	0.99
MLE-exp	0.01	0.53	0.71	0.82	0.95		0.27	0.71	0.83	0.89	0.97
MLE-iso	0.02	0.55	0.75	0.85	0.96		0.02	0.65	0.82	0.9	0.98

Table EC.10 Performance when methods are evaluated using heterogeneous linking functions with $\eta_v = 5$ and $b = 0.26$.