Technical Appendix: Solution Algorithm

The utility can be decomposed into average ($\delta_{jt}$), idiosyncratic ($\eta_{ijt}$), and random elements ($\varepsilon_{ijt}$):

$$U_{ijt} = \delta_{jt} + \eta_{ijt} + \varepsilon_{ijt},$$

where:

$$\delta_{jt} = -p_{jt}\bar{\alpha} + x_{jt}\bar{\beta} + [\gamma(A_{jt}^a)^\gamma, (A_{jt}^c)^\gamma, (A_{jt}^v)^\gamma]^{\prime}\mu + \xi_{b} + \xi_{t} + \Delta \xi_{jt},$$

$$\eta_{ijt} = -p_{jt} \left[ \Pi_{a}D_{i} + \Sigma_{a}\nu_{ia} \right] + x_{jt} \left[ \Pi_{b}D_{i} + \Sigma_{b}\nu_{ib} \right] + [(A_{jt}^a)^\gamma, (A_{jt}^c)^\gamma, (A_{jt}^v)^\gamma]^{\prime} \left[ \Pi_{a}D_{i} + \Sigma_{a}\nu_{ia} \right],$$

$$\varepsilon_{ijt} \sim \text{Extreme Value}$$

Following the literature, the specification of demand system is completed by normalizing the utility of an outside good (the choice not to consume any of the products) to zero. In this model framework an individual is defined as a vector of demographics and product specific shocks, ($D_{i}, \nu_{i}; \xi_{0}, \ldots, \xi_{J}$). Formally, market shares are given by:

$$s_{jt} = \int_{\nu} \int_{D} s_{ijt} dP_{D}(D) dP_{\nu}(\nu) = \int_{\nu} \int_{D} \left[ \frac{e_{ijt}^{\delta_{jt}} + \eta_{ijt}}{\sum_{k=1}^{N} e_{ijt}^{\delta_{jt}} + \eta_{ijt}} \right] dP_{D}(D) dP_{\nu}(\nu)$$

Isolating the unobservable quality component $\Delta \xi_{jt}$ is an essential step in this estimation procedure. The problem requires finding the values $\tilde{\delta}_{t} = \{ \tilde{\delta}_{1t}, \ldots, \tilde{\delta}_{Jt} \}$ of the average characteristics that equate the predicted and observed market shares of the $J_{t}$ products at time $t$. As discussed next, the objective function for this problem is a standard nonlinear least-squares problem:

$$\min_{\delta_{t}}: \frac{1}{2} \varepsilon \left( \tilde{\delta}_{t} \right)^{\prime} \varepsilon \left( \tilde{\delta}_{t} \right),$$

where

$$\varepsilon \left( \tilde{\delta}_{t} \right) = \begin{bmatrix} \bar{s}_{1t} - s_{1t}^{\text{obs}} \\ \vdots \\ \bar{s}_{Jt} - s_{Jt}^{\text{obs}} \end{bmatrix}$$

We use the notation $\bar{s}_{jt}$ to emphasize that these predicted market shares are derived from the average characteristics $\tilde{\delta}_{jt}$, which are not the same as the characteristics $\delta_{jt}$ in the utility specification. This algorithm determines all $\delta_{t}$ simultaneously by solving systems of nonlinear equations (e.g., Nocedal and Wright 2006; Nevo 2001). Because the $s_{jt}$ are smooth in $\delta_{t}$, this method is efficient in practice.
The integral in $s_{jt}$ is approximated by simulation with $R$ draws from the joint density of $d_t$ and $\nu_t$

$$s_{jt} \approx \frac{1}{R} \sum_{r=1}^{R} \left[ \exp \left( \delta_{jt} + \eta_{rjt} \right) \right] \left[ \frac{1}{1 + \sum_{k=1}^{J_t} \exp \left( \delta_{kt} + \eta_{rkt} \right)} \right]$$

The observable characteristics $d_t$ of these “simulated consumers” (age, income, and education) come from the U.S. Current Population Survey (CPS).\(^1\) The demographic vector consists of either 1 or 0 if consumer falls within that specific demographic segment or not, respectively. Because each element of the vector of demographics is a categorical variable, and if there are $N$ groups within a particular demographic (e.g., three age groups), only $N - 1$ interactions can be included in the model. Therefore, the mean parameter can be interpreted as already containing the $N$th group, which might be avoided if the demographic variables were continuous. However, to be consistent with the Mediamark data on micro moments, we need this grouping. Each consumer therefore is defined by $3 - 1 = 2$ age dummies, $2 - 1 = 1$ income dummies, and $2 - 1 = 1$ education attainment dummies, for a total of $2 + 1 + 1 = 4$ elements of demographic vector. The unobservable characteristics $\nu_t$ are drawn from a normal distribution with zero mean and unitary variance.

An exact or approximate representation of the Jacobian matrix of $\varepsilon(\tilde{d}_t)$ is required in a trust-region implementation of the nonlinear equations method. The exact Jacobian matrix for this problem contains derivatives of the form:

$$\frac{\partial s_{jt}}{\partial \delta_{mt}} = \begin{cases} E_{d_t, \nu_t} \left[ \frac{\exp \left( \delta_{jt} + \eta_{ijt} \right)}{1 + \sum_{k=1}^{J_t} \exp \left( \delta_{kt} + \eta_{ikt} \right)} \right] \cdot \left( 1 - \frac{\exp \left( \delta_{jt} + \eta_{ijt} \right)}{1 + \sum_{k=1}^{J_t} \exp \left( \delta_{kt} + \eta_{ikt} \right)} \right) & \text{if } j = m \\ -E_{d_t, \nu_t} \left[ \frac{\exp \left( \delta_{jt} + \eta_{ijt} \right)}{1 + \sum_{k=1}^{J_t} \exp \left( \delta_{kt} + \eta_{ikt} \right)} \right] \cdot \frac{\exp \left( \delta_{jt} + \eta_{ijt} \right)}{1 + \sum_{k=1}^{J_t} \exp \left( \delta_{kt} + \eta_{ikt} \right)} & \text{else} \end{cases}$$

The elements within the expectations are individual brand shares, so a simulated approximation can be easily accumulated contemporaneously with the computation of the $s_{jt}$.

The overall problem thus is to employ GMM to optimize the model with respect to the parameters:

$$\theta = \left\{ \alpha, \Pi_\alpha, \Sigma_\alpha, \beta, \Pi_\beta, \Sigma_\beta, \mu, \Pi_\mu, \Sigma_\mu, \left\{ \xi_b \right\}_{b=1}^{B}, \left\{ \xi_t \right\}_{t=1}^{T} \right\}$$

\(^1\)Instead of assuming parametric forms for the distribution of demographics, the simulation is performed using the empirical non-parametric distribution of real individuals from the March CPS, which implicitly accounts for the correlations in demographics and can be interpreted as a nonparametric estimate of the joint distribution of demographics.
Macro Moments

The theoretical macro moments are of the form $E[z\Delta\xi] = 0$, where $z$ is a vector of instruments. The variance of the macro moments is: $E[(z\Delta\xi)(z\Delta\xi)']$. Given $\delta_t$, a value $\Delta\xi_{jt}^{(r)}$ for a simulated consumer $r$ can be determined from the utility specification and the value of $\delta_{jt}$:

$$
\Delta\xi_{jt}^{(r)} = \delta_{jt} - (-p_{jt} [\alpha + \Pi_a D_i + \Sigma_a \nu_i a] + x_{jt} [\beta + \Pi_{\beta} D_i + \Sigma_{\beta} \nu_{i\beta}] + 
+ [(A^n_{jt})^\gamma, (A^{c}_{jt})^\gamma, (A^{oc}_{jt})^\gamma]^T \Pi + \Pi \mu D_i + \Sigma \mu \nu_i \mu] + \xi_b + \xi_t)
$$

These simulated values give rise to the following sample analog to macro moments:

$$
G_1 \equiv \frac{1}{T} \sum_{t=1}^{T} G_{1t} = \frac{1}{T} \sum_{t=1}^{T} \left[ \frac{1}{J_t R} \sum_{j=1}^{J_t} \sum_{r=1}^{R} z_{jt} \Delta\xi_{jt}^{(r)} \right] = 0.
$$

In addition, the analog to the sample variance of the macro moments is

$$
\Omega_1 = \frac{1}{T} \left[ \frac{1}{T-1} \sum_{t=1}^{T} G_{1t} G_{1t}' \right]
$$

(Greene 2004). The relevant dimension of sampling variability is time; the consumer-product variation represents simulation variability that does not affect the asymptotic properties of these estimates.

Micro Moments

The micro moments match predicted and observed market shares for brands within demographic groups. For example, the random coefficient on comparative advertising for a consumer who’s age falls within 18-35 age category, can be rewritten in the following way:

$$
\mu^c(\psi(A^c_j)) = [\bar{\mu}^c + \pi_1 D_i (18 < i's\ age < 35)] + \sigma \nu_i \] \psi(A^c_j)$$

where $\bar{\mu}^c$ is a mean valuation for comparative advertising, $\pi$ is an incremental valuation for a comparative advertising for an individual who is 18-35 years old and $\sigma$ measures how individual valuation on comparative advertising is affected by unobserved consumer characteristics. The time scale of the micro moments is six months, whereas that for the macro moments is one month. Thus, we temporally index the micro moments by the starting month of the six-month block, such that $t' = 6, 12, 18, 24, \ldots$ denotes the time points of interest with respect to the micro moments.\(^2\)

\(^2\)Data by demographics are unavailable for the first four periods of the sample.
Since during that time any consumer might have consumed more than one brand, we calculate
the index of relative consumption within a demographic group which necessitates the
calculation of ratios of conditional over unconditional brand consumption means.

The unconditional probability of purchasing brand $b$ at time $t$ is:

$$s_{bt} = \sum_{j_b=1}^{J_{bt}} s_{j_b t},$$

where $s_{j_b t}$ is the brand share for product $j_b$ at time $t$ (as defined by the logit specification),
and $J_{bt}$ is the total number of products corresponding to brand $b$ at time $t$. Similarly, the
probability of demographic $d^*$ purchasing brand $b$ at time $t$ is:

$$s_{d^* bt} = \sum_{j_b=1}^{J_{bt}} s_{j_b t}^*,$$

Using the same method we used for the macro moments, we can approximate the brand
shares using simulated consumers. In practice, this approximation means averaging brand
shares from $R$ simulated consumers to obtain the unconditional probability $s_{bt}$, and averaging
the brand shares from a subset $R_{d^*}$ (corresponding to simulated consumers of demographic
$d^*$) to obtain the conditional probability $s_{d^* bt}$.

The micro moment corresponding to the purchase of brand $b$ by demographic $d^*$ at time $t'$ takes
the form:

$$E \left( \frac{s_{d^* bt}}{s_{bt}} - y_{d^* bt} \right) = 0,$$

where $y_{d^* bt}$ denotes the observed brand share ratio. The variance of this moment is:

$$E \left[ \left( \frac{s_{d^* bt}}{s_{bt}} \right)^2 \right] - \left( y_{d^* bt} \right)^2.$$
For convenience, we order all micro moments by (a) the time block, then by (b) demographics within a block, and finally by (c) brands within a demographic. In addition, we define \( g_{2t} \) as the block-level vector that contains elements of the form \( \frac{s_{it}}{s_{it}} \) in the proper demographic-brand ordering. Let \( y_{t'} \) denote the corresponding vector of observed brand shares in the same ordering. With these definitions, the sample micro moments within each block can be written as:

\[
G_{2t'} = \frac{1}{\tau} \sum_{t=0}^{\tau-1} g_{2t'+t} - y_{t'}
\]

and the sample variance is:

\[
\Omega_{2t'} = \frac{1}{\tau} \left[ \frac{1}{\tau-1} \sum_{t=0}^{\tau-1} g_{2t'+t} g_{2t'+t}^T - y_{t'} y_{t'}^T \right]
\]

Using this notation, we define

\[
G_2 = \begin{bmatrix}
G_{2,t'=6} \\
G_{2,t'=12} \\
G_{2,t'=18} \\
\vdots
\end{bmatrix}
\]

and

\[
\Omega_2 = \begin{bmatrix}
\Omega_{2,t'=6} & \Omega_{2,t'=12} & \Omega_{2,t'=18} & \cdots
\end{bmatrix}
\]

This representation is possible because, by construction, a time-\( t \) observation applies to one and only one time block of duration \( \tau = 6 \).

**Estimation Steps**

The estimation proceeds in the following steps:

- **Step 1.** Find a consistent but inefficient estimate \( \tilde{\theta} \) using unweighted GMM. (Simulated annealing run #1),

- **Step 2.** Construct an optimal weighting matrix using \( \tilde{\theta} \),

- **Step 3.** Find a consistent and efficient estimate \( \hat{\theta} \) using weighted GMM. (Simulated annealing run #2).
For Step 1, the GMM minimization problem is

\[ \min_{\theta} : \begin{bmatrix} G_1(\theta) \\ G_2(\theta) \end{bmatrix}' \begin{bmatrix} G_1(\theta) \\ G_2(\theta) \end{bmatrix} \]

A parallel simulated-annealing algorithm is used to find a minimizer of this function. In the parallelization scheme, we use a total of \( T = 58 \) nodes, and each node is assigned the data for exactly one of the time period. The task of node \( t \) is to compute \( G_{1t}(\theta) \) and \( G_{2t}(\theta) \). The determination of steps in \( \theta \)-space is controlled by a master node that aggregates the \( G_{1t} \) and \( G_{2t} \) contributions from all nodes.

The estimate \( \tilde{\theta} \) found in Step 1 is likely to be inefficient, though it is consistent.

Step 2 constructs the optimal weighting matrices using \( \tilde{\theta} \) from the unweighted problem. The optimal weighting matrices are exactly the sample covariance matrices \( \Omega_1(\tilde{\theta}) \) and \( \Omega_2(\tilde{\theta}) \).

For Step 3, the GMM minimization problem is

\[ \min_{\theta} : \begin{bmatrix} G_1(\theta) \\ G_2(\theta) \end{bmatrix}' \begin{bmatrix} \Omega_1(\tilde{\theta}) & 0 \\ 0 & \Omega_2(\tilde{\theta}) \end{bmatrix}^{-1} \begin{bmatrix} G_1(\theta) \\ G_2(\theta) \end{bmatrix} , \]

where \( \Omega \) is a feasible estimate of the optimal weighting matrix.

The minimizer \( \hat{\theta} \) is obtained using the same parallel simulated-annealing algorithm used for Step 1. According to various feasible general least squares (GLS) theorems (e.g., Greene 2004), a \( \hat{\theta} \) obtained in this three-step process is both consistent and efficient.

Finally, the variance of \( \hat{\theta} \) is

\[ \text{Var}(\hat{\theta}) = \left[ J_1(\hat{\theta})' \Omega_1^{-1}(\hat{\theta}) J_1(\hat{\theta}) \right]^{-1} + \left[ J_2(\hat{\theta})' \Omega_2^{-1}(\hat{\theta}) J_2(\hat{\theta}) \right]^{-1} \]

where \( J_1(\theta) \) and \( J_2(\theta) \) denote the Jacobian matrices of \( G_1(\theta) \) and \( G_2(\theta) \) (again, see chapter 11 of Greene 2004).

In Table 15 we provide the complete set of model parameters and their dimensions.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Random price coefficient</td>
<td>$[1 \times 1]$</td>
</tr>
<tr>
<td>$\Pi_\alpha$</td>
<td>Observed heterogeneity for $p$</td>
<td>$[1 \times d]$</td>
</tr>
<tr>
<td>$\Sigma_\alpha$</td>
<td>Unobserved heterogeneity for $p$</td>
<td>$[1 \times 1]$</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>Random characteristics coefficients</td>
<td>$[k \times 1]$</td>
</tr>
<tr>
<td>$\Pi_\beta$</td>
<td>Observed heterogeneity for $X_1$</td>
<td>$[k \times d]$</td>
</tr>
<tr>
<td>$\Sigma_\beta$</td>
<td>Unobserved heterogeneity for $X_1$</td>
<td>$[k \times k]$, diagonal</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>Non-random (characteristics) coefficients</td>
<td>$[1 \times 1]$</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Random Ads coefficients</td>
<td>$[h \times 1]$</td>
</tr>
<tr>
<td>$\Pi_\mu$</td>
<td>Observed heterogeneity for Ads</td>
<td>$[h \times d]$</td>
</tr>
<tr>
<td>$\Sigma_\mu$</td>
<td>Unobserved heterogeneity for Ads</td>
<td>$[h \times h]$, diagonal</td>
</tr>
<tr>
<td>$\nu_{ij\beta}$, $\nu_{ij\mu}$</td>
<td>Random draws from N(0,1)</td>
<td>$[k \times 1], [h \times 1]$</td>
</tr>
<tr>
<td>$D_i$</td>
<td>Consumer (defined by demographics)</td>
<td>$[d \times 1]$</td>
</tr>
<tr>
<td>$\Delta \xi$</td>
<td>Unobserved quality</td>
<td>$[J \times 1]$</td>
</tr>
<tr>
<td>$Z$</td>
<td>Instruments</td>
<td>$[M \times J]$</td>
</tr>
<tr>
<td>$G_1$</td>
<td>Macro moment conditions</td>
<td>$[M \times 1]$</td>
</tr>
<tr>
<td>$G_2$</td>
<td>Micro moment conditions</td>
<td>$[t \times d \times B' \times 1]$</td>
</tr>
<tr>
<td>$\Omega_1$</td>
<td>Estimate of optimal weighing matrix for $G_1$</td>
<td>$[Z \times Z]$</td>
</tr>
<tr>
<td>$\Omega_2$</td>
<td>Estimate of optimal weighing matrix for $G_2$</td>
<td>$[t \times d \times B' \times t \times d \times B']$</td>
</tr>
<tr>
<td>$W(\theta)$</td>
<td>GMM function</td>
<td>$[1 \times 1]$</td>
</tr>
</tbody>
</table>

$N = 17,144$ Observations  
$J = 392$ Products  
$R = 250$ Simulated consumers  
$B = 10$ 6 brands + 4 types of generics  
$B' = 7$ 6 brands + 1 (combined generics)  
$T = 58$ Time periods  
$M = 59$ Instruments  
$d = 4$ Demographic points  
$k = 4$ Variables in $X_1$  
$l = 10$ Variables in $X_2$  
$h = 3$ Variables in $A$ ($A_{nc}, A_c, A_{oc}$)  
$t = 9$ 6 month blocks for micro moments

### Elasticity Approximation

We presented advertising elasticities in the text. The own- and cross- price elasticities are computed in a usual way:

$$
\frac{\partial s_{jt} p_{kt}}{\partial p_{kt} s_{jt}} = \begin{cases} 
-\frac{\nu_{jt}}{s_{jt}} \int \alpha_i s_{ijt} (1 - s_{ijt}) dp_D(D) dp_\nu(\nu) & \text{if } j = k \\
\frac{\nu_{jt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dp_D(D) dp_\nu(\nu) & \text{else} 
\end{cases} 
$$

(1)
The integrals in the advertising and price elasticity equations can not be computed analytically, therefore we use the computational approximation. In this case, \( P_D(D) \) is an empirical distribution of individuals (\( R \) draws from CPS data) and \( \nu \sim N(0,1) \). Then, the integral of individual market shares, \( s_{ijt} \), is replaced with simulated market shares (\( \bar{s}_{ijt} \)) by taking \( R \) draws of consumers.

References


Appendix A: Product Characteristics

Medical Properties

We reviewed medical journal articles to establish or rank the medical properties of the analyzed active ingredients. Most medical articles compare only two or three active ingredients. If article X said that drug $A$ is more efficient than drug $B$ ($A \succ B$) and article Y said that drug $B$ is more efficient than $C$ ($B \succ C$), we conclude by transitivity that $A$ is more efficient than $B$ and $C$ ($A \succ B \succ C$). The assembled information on medical properties is plotted in Figure 2.

Figure 2. Location of Active Ingredients in the Characteristics Space
The number-needed-to-treat (NNT) is computed with respect to two treatments A and B, such that A is typically a drug and B a placebo. If the probabilities \( P_A \) and \( P_B \) for treatments A and B, respectively, are known, the NNT can be computed as:

\[
NNT = \frac{1}{P_B - P_A}
\]

The NNT for a given therapy is simply the reciprocal of the absolute risk reduction (ARR = \( P_B - P_A \)) for that treatment. For example, in the hypothetical migraine study, risk decreased from \( P_B = 0.30 \) without treatment with drug M to \( P_A = 0.05 \) with treatment with drug M, for a relative risk of 0.17 \((0.05/0.3)\), a relative risk reduction of 0.83 \(((0.3-0.05)/0.3)\), and an absolute risk reduction of 0.25 \((0.3-0.05)\), the NNT would be \(1/0.25\), or 4. In concrete clinical terms, an NNT of 4 means that four patients need treatment with drug M to prevent a migraine from recurring in one patient. Typically, the lower the NNT, the more potent and efficient the treatment is.

Cardiovascular Risk (CV) and Gastrointestinal Risk (GI) are expressed as relative risk (RR) measures. Relative risk is the risk of an event (e.g., developing a disease) relative to exposure. Relative risk is the ratio of the probability of the event \( E \) occurring in the exposed group versus in the control (unexposed) group:

\[
RR = \frac{Pr(E | \text{treatment})}{Pr(R | \text{control})}
\]

In clinical trial data, RR serves to compare the risk of developing a disease in people who have not received a new medical treatment (or receive a placebo) versus that in people who receive an established (standard of care) treatment. For the GI and CV relative risk numbers we employ, RR helps compare the risk of developing a side effect in people who receive a drug compared with that for people who do not. Thus, a CV RR of 1.44 means that CV problems arise with 44% higher likelihood in people using the drug (versus a placebo).

**Other Characteristics**

Some marketed pain relievers contain additional active ingredients that help with specific pain relief (e.g., nighttime pain relief formulas) or affect the efficiency of pain relief.

*Caffeine.* Analgesic active ingredients combining caffeine with aspirin, acetaminophen, or both are available as OTC drugs (e.g., Excedrin for headache and migraine, Bayer for body aches). Caffeine is an analgesic adjuvant that enhances the analgesic effects of aspirin, acetaminophen, and ibuprofen; it is ineffective when used alone.
Aspirin/acetaminophen combinations. Aspirin and acetaminophen combination enhance the efficiency of pain relief by adding active ingredients that produce the same effect by different channels. These combinations should have fewer side effects than aspirin alone. Excedrin, Midol and Pamprin use this mix.

Other formulas. Night formulae (e.g., Tylenol PM) contain an additional ingredient, diphenhydramine HCl, that is a sleep aid. Drugs marketed for menstrual purposes (e.g., Midol, Pamprin) with acetaminophen-based formulae often contain diuretics such as pamabrom and pyrilamine maleate.

Table 14 lists the advertised brands and products, together with their active ingredients, recommended dosage, and maximum number of pills allowed within 24 hours.

<table>
<thead>
<tr>
<th>Brand</th>
<th>Product</th>
<th>Active Ingredients</th>
<th>Dosage</th>
<th>24h Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advil</td>
<td>Advil Liquidgels</td>
<td>Sol. Ibuprofen 200mg</td>
<td>1 every 4 to 6 h</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Advil</td>
<td>Ibuprofen 200mg</td>
<td>1 every 4 to 6 h</td>
<td>6</td>
</tr>
<tr>
<td>Aleve</td>
<td>Aleve</td>
<td>Naproxen Sodium 220 mg</td>
<td>1 every 8 to 12 h</td>
<td>3</td>
</tr>
<tr>
<td>Bayer</td>
<td>Bayer</td>
<td>Aspirin 325 mg</td>
<td>2 every 4 to 6 h</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Bayer Back and Body</td>
<td>Aspirin 500 mg, Caffeine 32 mg</td>
<td>2 every 6 h</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Bayer Extra Strength</td>
<td>Aspirin 500 mg</td>
<td>2 every 4 to 6 h</td>
<td>8</td>
</tr>
<tr>
<td>Excedrin</td>
<td>Excedrin</td>
<td>Acetaminophen 250 mg,</td>
<td>2 every 6 h</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Excedrin Tension Headache</td>
<td>Acetaminophen 500 mg, Caffeine 65 mg</td>
<td>2 every 6 h</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Excedrin Migraine</td>
<td>Acetaminophen 250 mg, Aspirin 250 mg, Caffeine 65 mg</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Excedrin Sinus Headache</td>
<td>Acetaminophen 325 mg, Phenylephrine HCl 5 mg</td>
<td>2 every 4 h</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Excedrin PM</td>
<td>Acetaminophen 500 mg, Diphenhydramine HCl 38 mg</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Motrin</td>
<td>Motrin IB</td>
<td>Ibuprofen, 200mg</td>
<td>1 every 4 to 6 h</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Children’s Motrin</td>
<td>Ibuprofen, 100mg</td>
<td>1 every 6 h</td>
<td>4</td>
</tr>
<tr>
<td>Tylenol</td>
<td>Tylenol Arthritis</td>
<td>Acetaminophen 650 mg</td>
<td>2 every 8 h</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Tylenol Regular</td>
<td>Acetaminophen 325 mg</td>
<td>2 every 4 to 6 h</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Tylenol Extra Strength</td>
<td>Acetaminophen 500 mg</td>
<td>2 every 4 to 6 h</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Tylenol PM</td>
<td>Acetaminophen 500 mg, Diphenhydramine HCl 25 mg</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Tylenol Rapid Release</td>
<td>Acetaminophen 500 mg</td>
<td>2 every 4 to 6 h</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Tylenol 8 hour</td>
<td>Acetaminophen 650 mg</td>
<td>2 every 8 h</td>
<td>6</td>
</tr>
</tbody>
</table>
References


